

Result	Score	Match	Length	DB	ID	Description
Query						

US-09-882-227-180

Alignment Scores:

Pred. No.: 1 59e-110 Length: 210
 Score: 1072.00 Matches: 207
 Percent Similarity: 99.05% Conservatives: 1
 Best Local Similarity: 98.57% Mismatches: 2
 Query Match: 68.76% Indels: 0
 DB: 10 Gaps: 0

US-09-647-661-1 (1-900) x US-09-882-227-180 (1-210)

QY 151 ATGAAATTTTGGATCAGGAAAAAGAGCACTATTAAACGAGCGCCATTCTTCAAG 210
 Db 1 MetLysPheLeuAspGlnGluYsArgArgGlnLeuLeuAsnGluArgHisSerCysLys 20
 QY 211 ATGTTTCACAGCCATTATGAGTTTCTAGTGAAGAATTAGCAAAATCGTGAATCGCC 270
 Db 21 MetPheAspSerHisTyrGluPheSerSerThrGluLeuGluGluLeuAlaGluLeuAla 40
 QY 271 AGGCTATCGCAAGCTCTTAAACACGAGCCATGCGATTTTGTGATGTTACTATAAG 330
 Db 41 ArgLeuSerProSerSerTyrAsnThrGlnProTrpHisPheValMetValThrAspLys 60
 QY 331 GATTTAARAAACAAATTCAGTCACAGCTACTTTAATCAAGAAATGATTAAACCGCT 390
 Db 61 AspLeuLysGlnIleAlaAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
 QY 391 TCACGCTTAATGTTGATGCTCTTTAAGACCTAGCGAGTTGTACACAGCGCCATTAC 450
 Db 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTyr 100
 QY 451 ATGCAAAACCTTTACCGGAGTCTTAAAGTAGAGTGTATGAGTGTATGCTCTTCTCAATG 510
 Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
 QY 511 CTTCGCGTGAAGTCAACACGAGCATGCAAGAGTATAGAAAGCTATATTAGAGCAATGC 570
 Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGluGlnCys 140
 QY 571 TATATCGCTGTGGGCAAAATTCATGAGCGGTGAGCTTAATGGATGATAGTTCATTT 630
 Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
 QY 631 ATTGAGGCTTTGATCTCTTTAAAGTGGGTGAAGTTTGAAGAGCGCTATCAATAAGCT 690
 Db 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluGluArgIleAsnLysPro 180
 QY 691 AAAATCGCATGCTTGATCGCTTTGGCAAGAGGTGGCAGAGCGCCAAATCAAGA 750
 Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
 QY 751 AAATCAAAAGTTGATGCTTACTTTGTTG 780
 Db 201 LysSerLysValAspAlaIleThrIrpueu 210

RESULT 2

US-10-364-397-29
 ; Sequence 29, Application US/10364397
 ; Publication No. US20040014191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Minton, Nigel
 ; APPLICANT: Anlezark, Gail
 ; APPLICANT: Vaughan, Thomas
 ; TITLE OF INVENTION: Microreductase Enzymes
 ; FILE REFERENCE: 1581/0650001
 ; CURRENT APPLICATION NUMBER: US/10/364,397
 ; CURRENT FILING DATE: 2003-02-12
 ; PRIOR APPLICATION NUMBER: 09/913,068
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00431
 ; PRIOR FILING DATE: 2000-02-10
 ; PRIOR APPLICATION NUMBER: UK 9903019.9

; PRIOR FILING DATE: 1999-02-10
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 29
 ; LENGTH: 210
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 US-10-364-397-29

Alignment Scores:

Pred. No.: 5 73e-110 Length: 210
 Score: 1067.00 Matches: 206
 Percent Similarity: 99.05% Conservatives: 2
 Best Local Similarity: 98.10% Mismatches: 2
 Query Match: 68.44% Indels: 0
 DB: 15 Gaps: 0

US-09-647-661-1 (1-900) x US-10-364-397-29 (1-210)

QY 151 ATGAAATTTTGGATCAGGAAAAAGAGCACTATTAAACGAGCGCCATTCTTCAAG 210
 Db 1 MetLysPheLeuAspGlnGluYsArgArgGlnLeuLeuAsnGluArgHisSerCysLys 20
 QY 211 ATGTTTCACAGCCATTATGAGTTTCTAGTGAAGAATTAGCAAAATCGTGAATCGCC 270
 Db 21 MetPheAspSerHisTyrGluPheSerSerThrGluLeuGluGluLeuAlaGluLeuAla 40
 QY 271 AGGCTATCGCAAGCTCTTAAACACGAGCCATGCGATTTTGTGATGTTACTATAAG 330
 Db 41 ArgLeuSerProSerSerTyrAsnThrGlnProTrpHisPheValMetValThrAsnLys 60
 QY 331 GATTTAARAAACAAATTCAGTCACAGCTACTTTAATCAAGAAATGATTAAACCGCT 390
 Db 61 AspLeuLysGlnIleAlaAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
 QY 391 TCACGCTTAATGTTGATGCTCTTTAAGACCTAGCGAGTTGTACACAGCGCCATTAC 450
 Db 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProThrGlyHisTyr 100
 QY 451 ATGCAAAACCTTTACCGGAGTCTTAAAGTAGAGTGTATGAGTGTATGCTCTTCTCAATG 510
 Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
 QY 511 CTTCGCGTGAAGTCAACACGAGCATGCAAGAGTATAGAAAGCTATATTAGAGCAATGC 570
 Db 121 LeuGlyValArgPheAsnHisSerMetGlnLysLeuGluSerTyrIleLeuGluGlnCys 140
 QY 571 TATATCGCTGTGGGCAAAATTCATGAGCGGTGAGCTTAATGGATGATAGTTCATTT 630
 Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
 QY 631 ATTGAGGCTTTGATCTCTTTAAAGTGGGTGAAGTTTGAAGAGCGCTATCAATAAGCT 690
 Db 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluGluArgIleAsnLysPro 180
 QY 691 AAAATCGCATGCTTGATCGCTTTGGCAAGAGGTGGCAGAGCGCCAAATCAAGA 750
 Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
 QY 751 AAATCAAAAGTTGATGCTTACTTTGTTG 780
 Db 201 LysSerLysValAspAlaIleThrIrpueu 210

RESULT 3

US-10-335-977-6487
 ; Sequence 6487, Application US/10335977
 ; Publication No. US20040052799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DOUGLAS SMITH et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; RELATING TO HELICOBACTER PYLORI FOR
 ; DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 10031

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 6487:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...210
SEQUENCE DESCRIPTION: SEQ ID NO: 6487:
US-10-335-977-6487

Alignment Scores:
Pred. No.: 4.58e-107 Length: 210
Score: 1041.00 Matches: 200
Percent Similarity: 97.14% Conservative: 4
Best Local Similarity: 95.24% Mismatches: 6
Query Match: 66.77% Indels: 0
DB: 12 Gaps: 0

US-09-647-661-1 (1-900) x US-10-335-977-6487 (1-210)
QY 151 ATGAAATTTTGGATCAGGAAAAAGAGACAACATTTAAACGAGCGCCATTCTTGCAAG 210
DB 1 MetLysPheLeuAspHisGluLysArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTTGACGCCATTATGAGTTTCTAGTGAGAAATTAGAGAAATCGCTGAATCGCC 270
DB 21 MetPheAspSerHisTyrGluPheSerThrGluLeuGluLeuLeaGluLeaAla 40
QY 271 AGGTATCGCAAGCTCTTACACACGACGACCATTTGTCATGGTTACTATAAG 330
DB 41 ArgLeuSerProSerSerTyrAsnThrGlnProTrpHisPheValMetValThrAsnLys 60
QY 331 GATTAAAAAACAATTCGAGTCGACAGCTACTTTAATGAAGAAATGATTAAGCGGT 390
DB 61 AspLeuLysLysGlnIleAlaAlaHisSerTyrPheAsnGluLeuMetIleLysSerAla 80
QY 391 TCAGCGTTAATGGTGGTATCTCTTAAAGACCTAGCGAGTTGTACACACGCCATTAC 450
DB 81 SerAlaLeuMetValValCysSerLeuLysPheSerGlnLeuLeuProThrSerHisTyr 100
QY 451 ATGCAAAACCTTACCGGAGTCTTATAAGTTAGAGTATCCCTCTTTGCTCAATG 510
DB 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120

```

511 CTTGGCGTGAGATTCAACACGACGATCAAGACATTAGAAAGCTATATTTAGAGCAATGC 570
121 LeuGlyValArgPheAsnHisSerMetGlnLysLeuGluSerTyrIleLeuGluGlnCys 140
571 TATATCGCTGTGGGGCAAAATTTGTCATGGCGCTGAGCTTAATGGGATTGGATTGCAATT 630
141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
631 ATTGAGGCTTTGATCCCTTTAAAGTGGTCAAGTTTATAGAAAGCGGTATCAATAAGCCT 690
161 IleGlyGlyPheAspProLeuLysValGlyGlnIleLeuGluArgIleAsnLysPro 180
691 AAATCGCATGCTTCATCGCTTTGGGCAAGAGGCTGGCAGAGCGAGCCAAATCAAGA 750
181 LysIleValCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerLysLysSerArg 200
751 AAATCAAAAGTTGATGGGATTACTTGGTTG 780
201 LysSerLysValAspAlaIleThrIlePhe 210

RESULT 4
US-10-335-977-6486
; Sequence 6486, Application US/10335977
; Publication No. US20040052799A1
GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 6486:
SEQUENCE CHARACTERISTICS:
LENGTH: 134 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...134
SEQUENCE DESCRIPTION: SEQ ID NO: 6486:
US-10-335-977-6486
Alignment Scores:
Pred. No.: 6.47e-62 Length: 134

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 20, 2004, 19:19:48 ; Search time 67.5 Seconds

(without alignments)
8413.819 Million cell updates/sec

Title: US-09-647-661-1

Perfect score: 1559
Sequence: 1 tgcagattttacagacagc.....gggggtgttttcagcgcttc 900

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2.1/USPTO_epool_p/US09647661/runat_20042004.135348.2676/app_query.fasta_1.1095
-DB=SPTREMBL_25 -QPMT=fastan -SUFFIX=sept -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=sept -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09647661.QCGN.1.1.86 @runat_20042004.135348.2676 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_25:
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1087	69.7	210	2	O30585 helicobacte

2	1080	69.3	210	2	O8RJ56	O8RJ56 helicobacte
3	1077	69.1	210	2	O9FBD5	O9FBD5 helicobacte
4	1077	69.1	210	2	O9FBD8	O9FBD8 helicobacte
5	1077	69.1	210	2	O8VTN5	O8VTN5 helicobacte
6	1076	69.0	210	2	O9FBD0	O9FBD0 helicobacte
7	1074	68.9	210	2	O9FBD1	O9FBD1 helicobacte
8	1074	68.9	210	2	O9FBE7	O9FBE7 helicobacte
9	1074	68.9	210	2	O8RJ83	O8RJ83 helicobacte
10	1073	68.8	210	2	O8RJCA	O8RJCA helicobacte
11	1073	68.8	210	2	O9FBC9	O9FBC9 helicobacte
12	1073	68.8	210	2	O8VTP5	O8VTP5 helicobacte
13	1073	68.8	210	2	O8VTN9	O8VTN9 helicobacte
14	1072	68.8	210	2	O9FBD3	O9FBD3 helicobacte
15	1072	68.8	210	16	O25608	O25608 helicobacte
16	1069	68.6	210	2	O9FBD4	O9FBD4 helicobacte
17	1069	68.6	210	2	O8RSS2	O8RSS2 helicobacte
18	1068	68.5	210	2	O8VTP7	O8VTP7 helicobacte
19	1067	68.4	210	2	O8RJES	O8RJES helicobacte
20	1067	68.4	210	2	O8VTP1	O8VTP1 helicobacte
21	1067	68.4	210	2	O9FBE5	O9FBE5 helicobacte
22	1065	68.3	210	2	O8VTP2	O8VTP2 helicobacte
23	1065	68.3	210	2	O9FBE6	O9FBE6 helicobacte
24	1065	68.3	210	2	O9FBE1	O9FBE1 helicobacte
25	1065	68.3	210	2	O8VTN7	O8VTN7 helicobacte
26	1062	68.1	210	2	O8VTP6	O8VTP6 helicobacte
27	1059	67.9	210	2	O8VTP3	O8VTP3 helicobacte
28	1058	67.9	210	2	O9FBE3	O9FBE3 helicobacte
29	1058	67.9	210	2	O9FBD2	O9FBD2 helicobacte
30	1057	67.8	210	2	O8VTN6	O8VTN6 helicobacte
31	1057	67.8	210	2	O9FBE0	O9FBE0 helicobacte
32	1057	67.8	210	2	O9FBE4	O9FBE4 helicobacte
33	1057	67.8	210	2	O9FBE9	O9FBE9 helicobacte
34	1056	67.7	210	2	O9AHD6	O9AHD6 helicobacte
35	1056	67.7	210	2	O9AHF8	O9AHF8 helicobacte
36	1056	67.7	210	2	O9FBE8	O9FBE8 helicobacte
37	1054	67.6	210	2	O8RJAA	O8RJAA helicobacte
38	1052	67.5	210	2	O9FBD6	O9FBD6 helicobacte
39	1051.5	67.4	209	2	O8RJF5	O8RJF5 helicobacte
40	1049	67.3	210	2	O9RN84	O9RN84 helicobacte
41	1047	67.2	210	2	O9FBE2	O9FBE2 helicobacte
42	1047	67.2	210	2	O9FBE2	O9FBE2 helicobacte
43	1041	66.8	210	16	O92KP7	O92KP7 helicobacte
44	1040	66.7	210	2	O9FBD0	O9FBD0 helicobacte
45	1040	66.7	210	2	O9RN86	O9RN86 helicobacte

ALIGNMENTS

RESULT 1

ID	O30585	PRELIMINARY;	PRT;	210 AA.
AC	O30585			
DT	01-JAN-1998	(TREMBLrel. 05, Created)		
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)		
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)		
DE	NADPH-linked flavin nitroreductase.			
GN	RDXA.			
OS	Helicobacter pylori (Campylobacter pylori).			
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;			
OC	Helicobacteraceae; Helicobacter.			
OX	NCBI_TaxID=210;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=HP500;			
RX	MEDLINE=98283705; PubMed=9622362;			
RA	Goodwin A., Kersault D., Sisson G., Veldhuizen van Zanten S.J.,			
RA	Berg D.E., Hoffman P.S.;			
RT	"Metronidazole resistance in Helicobacter pylori is due to null			
RT	mutations in a gene (rdxa) that encodes an oxygen-insensitive NADPH			
RT	nitroreductase."			
RL	Mol. Microbiol. 28:383-393(1998).			
DR	EMBL; AF012552; AAC46349.1; "			
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.			

DR GO:0006118; P:electron transport; IEA.
DR InterPro: IPR002086; Aldehyde dehydr.
DR InterPro: IPR000415; Nitroreductase.
DR Pfam: PF00881; Nitroreductase; 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
SQ SEQUENCE 210 AA; 24123 MW; EF065C8BA3F0EBB1 CRC64;

Alignment Scores:
Pred. No.: 6,91e-101 Length: 210
Score: 1087.00 Matches: 210
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.72% Indels: 0
DB: 2 Gaps: 0

US-09-647-661-1 (1-900) x 030585 (1-210)

QY 151 ATGAAATTTTGGATCAGGAAAAAGAGACAACTATTAAACGAGCGCCATCTTGCAG 210
DB 1 MetLysPheLeuAspGlnGluYsArgGlnLeuLeuAenGluArgHisSerCysLys 20
QY 211 ATGTTTGACGACATATAGTTTCTAGTGAAGATTAGAGAAATCGCTGAATCGCC 270
DB 21 MetPheAspSerHisTyrGluPheSerSerGluGluLeuGluGluAlaGluAla 40
QY 271 AGGCTATCGCAAGCTCTTCAACACGAGCGATGGCATTTTGTGATGTTACTATAAG 330
DB 41 ArgLeuSerProSerSerTyrAsnThrGlnProTrpHisPheValMetValThrAsnLys 60
QY 331 GATTTAAAAACAAATTTGCGATGCGACAGCTACTTTAATGAAGAAATGATTAAGCGCT 390
DB 61 AspLeuLysLysGlnIleAlaValHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 351 TCAGCGTTAATGGTGTATGCTTTTAAGACTAGCGAGTTGTTACACAGCGCCATTAC 450
DB 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTyr 100
QY 451 ATGCAAAACCTTTACCGAGCTTTATAAGTTAGATGATCCCTTTCTTTCCTCAATG 510
DB 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
QY 511 CTGCGCTGAGATTTCAACACGAGCTGCAAGAGTTAGAACTATATTTAGAGCAATGC 570
DB 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGlnCys 140
QY 571 TATATCGCTGTGGGCAAAATTTGATGCTTTGGCAAGGGTGGCAGAGCGCAAAATCAAGA 630
DB 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGAGCGCTTTGATCCTTTAAAGTGGTGAAGTTTAGAGCGCTATCAATAAGCCT 690
DB 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluGluArgIleAsnLysPro 180
QY 691 AAAATCGCATGCTTGATGCTTTGGCAAGGGTGGCAGAGCGCGCAAAATCAAGA 750
DB 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAATCAAAAGTTGATGCGATTACTCGTTG 780
DB 201 LysSerLysValAspAlaIleThrTrpLeu 210

RESULT 2
ID Q8RJ56 PRELIMINARY; PRT; 210 AA.
AC Q8RJ56;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE RdxA protein.
GN RDXA
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.

OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=190A, and 190C;
RA Bereswill S., Krainick C., Herrmann L., Kist M.;
RT "Absence of rdxA Mutations in High-level Metronidazole Resistant
RT Clinical Isolates of Helicobacter pylori."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ305350; CAC83802.1; -;
DR EMBL; AJ305351; CAC83803.1; -;
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0016491; P:electron transport; IEA.
DR InterPro: IPR002086; Aldehyde dehydr.
DR InterPro: IPR000415; Nitroreductase.
DR Pfam: PF00881; Nitroreductase; 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
SQ SEQUENCE 210 AA; 24067 MW; 8581158BBB47A88D CRC64;

Alignment Scores:
Pred. No.: 3,51e-100 Length: 210
Score: 1080.00 Matches: 208
Percent Similarity: 99.52% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 1
Query Match: 69.28% Indels: 0
DB: 2 Gaps: 0

US-09-647-661-1 (1-900) x Q8RJ56 (1-210)

QY 151 ATGAAATTTTGGATCAGGAAAAAGAGACAACTATTAAACGAGCGCCATCTTGCAG 210
DB 1 MetLysPheLeuAspGlnGluYsArgGlnLeuLeuAenGluArgHisSerCysLys 20
QY 211 ATGTTTGACGACATATAGTTTCTAGTGAAGATTAGAGAAATCGCTGAATCGCC 270
DB 21 MetPheAspSerHisTyrGluPheSerSerGluGluLeuGluGluAlaGluAla 40
QY 271 AGGCTATCGCAAGCTCTTCAACACGAGCGATGGCATTTTGTGATGTTACTATAAG 330
DB 41 ArgLeuSerProSerSerTyrAsnThrGlnProTrpHisPheValMetValThrAsnLys 60
QY 331 GATTTAAAAACAAATTTGCGATGCGACAGCTACTTTAATGAAGAAATGATTAAGCGCT 390
DB 61 AspLeuLysLysGlnIleAlaValHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 391 TCAGCGTTAATGGTGTATGCTTTTAAGACTAGCGAGTTGTTACACAGCGCCATTAC 450
DB 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTyr 100
QY 451 ATGCAAAACCTTTACCGAGCTTTATAAGTTAGATGATCCCTTTCTTTCCTCAATG 510
DB 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
QY 511 CTGCGCTGAGATTTCAACACGAGCTGCAAGAGTTAGAACTATATTTAGAGCAATGC 570
DB 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGlnCys 140
QY 571 TATATCGCTGTGGGCAAAATTTGATGCTTTGGCAAGGGTGGCAGAGCGCGCAAAATCAAGA 630
DB 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGAGCGCTTTGATCCTTTAAAGTGGTGAAGTTTAGAGCGCTATCAATAAGCCT 690
DB 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluGluArgIleAsnLysPro 180
QY 691 AAAATCGCATGCTTGATGCTTTGGCAAGGGTGGCAGAGCGCGCAAAATCAAGA 750
DB 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAATCAAAAGTTGATGCGATTACTCGTTG 780
DB 201 LysSerLysValAspAlaIleThrTrpLeu 210

RESULT 3

Qy	151	ATGA	ATTTT	TGAT	CAG	AAAA	AAGA	CA	ACT	TTAA	CAG	CGCC	ATCT	TG	CAAG	210										
Db	1	Met	Lys	Phe	Leu	Asp	Gln	Glu	Lys	Arg	Gln	Leu	Leu	Asn	Glu	Arg	His	Ser	Cys	Lys	20					
Qy	211	ATG	TTT	G	CAG	CAT	TAT	GAG	TTT	TCT	AGT	GA	GAT	TAG	AAG	AAT	TAG	A	GA	AAT	CGC	TG	AAAT	CGCC	270	
Db	21	Met	C	Phe	Asp	Ser	His	Tyr	Glu	Phe	Ser	Thr	Glu	Leu	Glu	Glu	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Leu	40
Qy	271	AGC	TAT	CGC	CAAG	CTT	TAC	AAC	ACG	CAG	CGC	ATG	GC	ATT	TTT	GTG	ATG	GT	TACT	AT	TAAG	330				
Db	41	Arg	Leu	Ser	Pro	Ser	Ser	Tyr	Asn	Thr	Gln	Pro	Irp	His	Phe	Val	Met	Val	Thr	Asn	Lys	60				
Qy	331	GAT	TTA	AAAA	AA	CA	AAT	TCC	AGT	CG	AC	GAT	CTT	TAA	TG	A	GA	GA	AAT	TG	TTAA	AGCGT	390			
Db	61	Asp	Leu	Lys	Lys	Gln	Leu	Leu	Val	His	Ser	Tyr	Phe	Asn	Glu	Glu	Met	Leu	Lys	Ser	Ala	80				
Qy	391	TC	AGC	GT	AA	TG	TGT	GTG	CTT	TAA	GAC	TAG	CG	AGT	TGT	TAC	CAC	ACG	CCCA	TTAC	450					
Db	81	Ser	Ala	Leu	Met	Val	Val	Cys	Ser	Leu	Arg	Pro	Ser	Glu	Leu	Leu	Pro	His	Gly	His	Tyr	100				
Qy	451	ATG	CA	AAAA	CT	TTA	CCG	GAG	TCT	TAT	TA	AGT	TTA	GAT	GAT	CC	TCT	TTT	TG	CT	CA	AAATG	510			
Db	101	Met	Gln	Asn	Leu	Tyr	Pro	Glu	Ser	Tyr	Lys	Val	Arg	Val	Leu	Pro	Ser	Phe	Ala	Gln	Met	120				
Qy	511	CTT	GGC	GT	GAG	ATT	CA	AC	CAC	AGC	ATG	CA	AA	GAT	TAG	AA	GC	AT	TAT	TTT	TAG	AGCA	ATGC	570		
Db	121	Leu	Gly	Val	Arg	Phe	Asn	His	Ser	Met	Gln	Lys	Leu	Glu	Ser	Tyr	Leu	Leu	Glu	Gln	Cys	140				
Qy	571	TAT	AT	CGC	GT	CGG	CG	AA	TTT	TG	AT	CGG	CGC	TG	AGC	TTA	TG	GGG	ATT	TG	GAT	GT	CGAT	630		

Db	141	TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle	160
Qy	631	ATTGGAGGCTTTGATCCCTTTTAAAGTGGGTGAAGTTTTAGAAGAGCGGTATCAATAAGCCT	690
Db	161	IleGlyCysPheAspProLeuLysValGlyGluValLeuGluA-GilleAsnLysPro	180
Qy	691	AAATGCGATCGCTTGATCGCTTTGGGCAAGAGGGTGCAGAGCGACGACCAAAATCAAGA	750
Db	181	LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg	200
Qy	751	AAATCAAAAGTTGATGCGATTACTTGGTTG	780
Db	201	LysSerLysValAspAlaValThrTrpLeu	210
RESULT 4			
Q9FBD8 PRELIMINARY; PRT; 210 AA.			
AC	Q9FBD8	01-MAR-2001 (TrEMBLrel. 16, Created)	
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)	
DE	RdxA	(Oxygen-insensitive NADPH nitroreductase).	
GN	RDXA		
OS	Helicobacter pylori (Campylobacter pylori).		
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;		
OC	Helicobacteraceae; Helicobacter.		
OX	NCBI_TaxID=210;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=338A;		
RX	MEDLINE=20356933; PubMed=10898705;		
RA	Solca N.M., Bernasconi M.V., Piffaretti J.C.;		
RT	"Mechanism of metronidazole resistance in Helicobacter pylori:		
RT	comparison of the rdxA gene sequences in 30 strains.;"		
RL	Antimicrob. Agents Chemother. 44:2207-2210(2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=338A;		
RA	Maggi-Solca N., Piffaretti J.-C.;		
RL	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=313;		
RA	Marais A., Bilardi C., Cantet F., Megraud F.;		
RT	"Characterization of genes involved in the metronidazole resistance in		
RT	nine strains of Helicobacter pylori isolated before and after		
RT	irradiation treatment.;"		
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF180410; AAGO10691; -		
DR	EMBL; AF323009; AAL37282.1; -		
DR	EMBL; AF323008; AAL37281.1; -		
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.		
DR	GO; GO:0006118; P:electron transport; IEA.		
DR	InterPro; IPR002086; Aldehyde dehydr.		
DR	InterPro; IPR000415; Nitroreductase.		
DR	Fiam; PF00861; Nitroreductase.1.		
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.		
SQ	SEQUENCE 210 AA; 24067 MW; 61B11D871026A081 CRC64;		
Alignment Scores:			
Pred. No.:	7.05e-100	Length:	210
Score:	1077.00	Matches:	208
Percent Similarity:	99.05%	Conservative:	0
Best Local Similarity:	99.05%	Mismatches:	2
Query Match:	69.08%	Indels:	0
DB:	2	Gaps:	0
US-09-647-661-1 (1-900) x Q9FBD8 (1-210)			
Qy	151	ATGAAATTTTGGATCAGGAAAGAGACAACTATTAAAGCGCCATCTTCGCAAG	210
Db	1	MetLysPheLeuAspGlnGluLysArgGlnLeuLeuAsnGluArgHisSerCysLys	20

Score: 1077.00 Matches: 207
Percent Similarity: 99.52% Conservative: 2
Best Local Similarity: 98.57% Mismatches: 1
Query Match: 69.08% Indels: 0
DB: Gaps: 0

US-09-647-661-1 (1-900) x Q8VTN5 (1-210)

QY 211 ATGTTTGACAGCCATTATGAGTTTCTAGTGAAGAAATAGAGAAATCGCTGAAATCGCC 270
DB 21 MetPheAspSerHisTyrGluPheSerSerThrGluLeuGluGluLeuAla 40
QY 271 AGCTATCGCCAAAGCTCTTACACACAGCAGCCATGCGATTGCTGCTTACTAATAAG 330
DB 41 ArgLeuSerProSerTyrAsnThrGlnProTyrPheValMetValThrAsnLys 60
QY 331 GATTAAAAAACAATTCAGTGCACAGCTACTTTAATGAAGAAATGATTAAGCGCT 390
DB 61 AspLeuLysLysGlnIleAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 391 TCAGCGTTAAGTGGTATGCTCTTTAAGACCTAGCAGTGTGTACACACAGCCCATAC 450
DB 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTyr 100
QY 451 ATGCACAACTTACCCGGAGTCTTATAAGTTAGAGTATGCTCTTTTGTCTCAATG 510
DB 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
QY 511 CTGGCGTGTGAGATTCAACACAGCATGCAAGAGTATAGAAAGCTATATTTAGAGCAATGC 570
DB 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGlnCys 140
QY 571 TATATCGCTGCGGCAATTTGCATGGCGTGCAGCTTAATGGGATGGATGCTGCAAT 630
DB 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGAGCGTTGTATCTCTTTAAAGTGGTGAAGTTTGAAGAGCGTATCAATAAGCCT 690
DB 161 IleGlyGlyPheAspProLeuLysValGlyValLeuGluGluArgIleAsnLysPro 180
QY 691 AAAATCGCATCTTGATCGCTTTGGGCAAGAGGTGGCAGAGCGCCAAAATCAAGA 750
DB 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAATCAAAAGTTGATGCGATTACTTGGTTG 780
DB 201 LysSerLysValAspAlaIleThrIleu 210

RESULT 5
Q8VTN5 PRELIMINARY; PRT; 210 AA.
ID Q8VTN5
AC Q8VTN5
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Oxygen-insensitive NADPH nitroreductase.
GN RDXA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=196;
RA Marais A., Cantet F., Megraud F.;
RT "Characterization of genes involved in the metronidazole resistance in
RT nine strains of Helicobacter pylori isolated before and after
RT eradication treatment."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF232018; AAL37291.1;
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002086; Aldehyde dehydr.
DR InterPro; IPR000415; Nitroreductase.
DR Pfam; PF00881; Nitroreductase; 1.
DR PROSITE; PS00070; ALDEHYDE DEHYDR. CYS; 1.
SQ SEQUENCE 210 AA; 24095 MW; E37E2986090FB52 CRC64;

Alignment Scores: 7.05e-100 Length: 210
Pred. No.:

Score: 1077.00 Matches: 207
Percent Similarity: 99.52% Conservative: 2
Best Local Similarity: 98.57% Mismatches: 1
Query Match: 69.08% Indels: 0
DB: Gaps: 0

US-09-647-661-1 (1-900) x Q8VTN5 (1-210)

QY 151 ATGAAATTTTGGATCGGAAAAAAGAGCAACTATTAACGAGCGCCATTCTTGTCAAG 210
DB 1 MetLysPheLeuAspGlnGluLysArgArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTTGACAGCCATTATGAGTTTCTAGTGAAGAAATAGAGAAATCGCTGAAATCGCC 270
DB 21 MetPheAspSerHisTyrGluPheSerSerGluGluLeuGluIleAlaGluIleAla 40
QY 271 AGCTATCGCCAAAGCTCTTACACACAGCAGCCATGCGATTGCTGCTTACTAATAAG 330
DB 41 ArgLeuSerProSerTyrAsnThrGlnProTyrPheValMetValThrAsnLys 60
QY 331 GATTAAAAAACAATTCAGTGCACAGCTACTTTAATGAAGAAATGATTAAGCGCT 390
DB 61 AspLeuLysLysGlnIleAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 391 TCAGCGTTAAGTGGTATGCTCTTTAAGACCTAGCAGTGTGTACACACAGCCCATAC 450
DB 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTyr 100
QY 451 ATGCACAACTTACCCGGAGTCTTATAAGTTAGAGTATGCTCTTTTGTCTCAATG 510
DB 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheSerGlnMet 120
QY 511 CTGGCGTGTGAGATTCAACACAGCATGCAAGAGTATAGAAAGCTATATTTAGAGCAATGC 570
DB 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGlnCys 140
QY 571 TATATCGCTGCGGCAATTTGCATGGCGTGCAGCTTAATGGGATGGATGCTGCAAT 630
DB 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGAGCGTTGTATCTCTTTAAAGTGGTGAAGTTTGAAGAGCGTATCAATAAGCCT 690
DB 161 IleGlyGlyPheAspProLeuLysValGlyValLeuGluLeuGluArgIleAsnLysPro 180
QY 691 AAAATCGCATCTTGATCGCTTTGGGCAAGAGGTGGCAGAGCGCCAAAATCAAGA 750
DB 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAATCAAAAGTTGATGCGATTACTTGGTTG 780
DB 201 LysSerLysValAspAlaIleThrIleu 210

RESULT 6
Q8VTN5 PRELIMINARY; PRT; 210 AA.
ID Q8VTN5
AC Q8VTN5
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE RDXA.
GN RDXA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=160A;
RX MEDLINE=20356933; PubMed=10898705;
RA Solca N.M., Bernasconi M.V., Piffaretti J.C.;
RT "Mechanism of metronidazole resistance in Helicobacter pylori:
RT comparison of the rdxA gene sequences in 30 strains."
RL Antimicrob. Agents Chemother. 44:2207-2210(2000).

```

RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN=160A;
RA Maggi-Solca N., Piffaretti J.-C.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF180398; AAG01057.1; -.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002086; Aldehyde dehydr.
DR InterPro; IPR000415; Nitroreductase.
DR Pfam; PF00881; Nitroreductase; 1.
DR PROSITE; PS00070; ALDEHYDE DEHYDR CYS; 1.
SQ SEQUENCE 210 AA; 24158 MW; 2DE724AFC2F89DBF CRC64;

Alignment Scores:
Pred. No.: 8,89e-100 Length: 210
Score: 1076.00 Matches: 208
Percent Similarity: 99.52% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 1
Query Match: 69.02% Indels: 0
DB: 2 Gaps: 0

US-09-647-661-1 (1-900) x Q9FBD1 (1-210)
QY 151 ATGAAATTTTGGATCAGGAAAAGAGACAACTTAAACGAGCGCATCTTGCAG 210
Db 1 MetLysPheLeuAspGlnGluSerGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTTGACAGCCATTATGATTTTCTAGTGAAGAATAGAGAAATCGCTGAATCGCC 270
Db 21 MetPheAspSerHisTyrGluPheSerSerGluLeuGluGluLeuAla 40
QY 271 AGGTATCGCCAGCTCTTACACAGCAGCCATTCGCTTGTGCTTACTATTAAG 330
Db 41 ArgLeuSerProSerTyrAsnThrGlnProTrpHisPheValMetValThrAsnLys 60
QY 331 GATTTAAAAAACAATTCAGTCCACAGCTACTTTAATGAAGAATGATTAAGCGCT 390
Db 61 AspLeuLysGlnLeuAlaValHisSerTyrPheAsnGluGluMetLysSerAla 80
QY 391 TCAGCGTTAATGGTATGCTCTTTAAGACCTAGCAGTGTGTACACAGCGCATAC 450
Db 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuGluGlnMet 120
QY 451 ATGCAAAACCTTTACCGGAGTCTTATAAGGTAGAGTATCCCTTTCTTTGCTCAAATG 510
Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
QY 511 CTGGCGTGAGATTCAACCAAGATTCAGAAAGCTATATATTTAGAGCAATGC 570
Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrLysSerAla 80
QY 571 TATATCGCTGTGGGCAATTTGCGAGGTGGAGTGTGATGGATTGGATAGTTGCATT 630
Db 141 TyrIleAlaValGlyGlnLeuLysValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGGAGCGCTTTGATCCCTTTAAAAAGTGGGTGAAGTTTAAAGAGCGGTATCAATAGCGCT 690
Db 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluGlnArgIleAsnLysPro 180
QY 691 AAAATCGCATCTTGATCGCTTTGGCAGAGGCTGGCAGAGCGAGCCCAAAATCAGA 750
Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAATCAAAAGTTGATGCCATTACTTGGTTG 780
Db 201 LysSerLysValAspAlaIleThrIlePleu 210

RESULT 7
Q9FBD1 PRELIMINARY; PRT; 210 AA.
AC Q9FBD1;
DT 01-MAR-2001 (TrEMBLrel. 16, C-created)

```

```

Db 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluGluArgIleAsnLysPro 180
QY 691 AAAATCGCATGCTGATCGCTTTGGCAAGAGGGTGGCAGAGCGCAAAAATCAAGA 750
Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAATCAAAAGTTCATCGGATCTACTGTTG 780
Db 201 LysSerLysValAspAlaIleThrTrpLeu 210

RESULT 8
Q9FBE7
ID Q9FBE7 PRELIMINARY; PRT; 210 AA.
AC Q9FBE7
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE RDXA.
GN RDXA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22A;
RX MEDLINE=20356933; PubMed=10898705;
RA Solca N.M., Bernasconi M.V., Piffaretti J.C.;
RT "Mechanism of metronidazole resistance in Helicobacter pylori:
RL Antimicrob. Agents Chemother. 44:2207-2210(2000)."
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=22A;
RA Maggì-Solca N., Piffaretti J.-C.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF180401; AG01060.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR02086; Aldehyde dehydr.
DR InterPro; IPR000415; Nitroreductase.
DR Pfam; PF00881; Nitroreductase; 1.
DR PROSITE; PS00070; ALDEHYDE DEHYDR. CYS; 1.
SQ SEQUENCE 210 AA; 24097 MW; 6B5BFF771CD4F7EC CRC64;

Alignment Scores:
Pred. No.: 1.42e-99 Length: 210
Score: 1074.00 Matches: 207
Percent Similarity: 99.05% Conservative: 1
Best Local Similarity: 98.57% Mismatches: 2
Query Match: 68.89% Indels: 0
DB: 2 Gaps: 0

US-09-647-661-1 (1-900) x Q9FBE7 (1-210)

QY 151 ATGAAATTTTGGATCAGGAAAAAGAACAACTATTAAACGAGCGCATCTTTCGAAG 210
Db 1 MetLysPheLeuAspGlnGluLysArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTGACAGCCATTAGAGTTTCTAGTGAAGAAATAGAAATCGCTGAATCGCC 270
Db 21 MetPheAspSerHisTyrGluPheSerThrGluLeuGluIleAlaGluIleAla 40
QY 271 AGGCTATCGCAAGCTCTTCAACACAGCGCATGCGATTTCGTGATGTTACTATAAG 330
Db 41 ArgLeuSerProSerSerTyrAsnThrGlnProThrPheValMetValThrAsnLys 60
QY 331 GATTATAAAAACAAATTCGAGTCACAGCTACTTTAATGAAGAAATGATTAACCGCT 390
Db 61 AspLeuLysLysGlnIleAlaValHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 391 TCAGCGTTAATGGTGGTATGCTCTTTAAGACCTAGCGAGTTGTACCACAGCGCCATTAC 450

```

```

Db 81 SerAlaLeuMetValValCysSerLeuLysProSerGluLeuLeuProHisGlyHisTyr 100
QY 451 ATGCAAAACCTTTACCGGAGTCTTTAAGGTTAGAGTGATCCCTCTCTTTTGTCTCAAATG 510
Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheThrGlnMet 120
QY 511 CTTGGCGTGAGATTCAACACACAGCATGCAAGATTAGAAAGCTATATTTAGAGCAATGC 570
Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGluGlnCys 140
QY 571 TATATCGCTGTGGGCAAAATTTGCATGGCGTGGAGTTAATGGATGGATGATTCATT 630
Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGGAGGCTTTCATCCTTTAAAGTGGGTGAAGTTTGAAGAGCGTATCAATAAGCCT 690
Db 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluGluArgIleAsnLysPro 180
QY 691 AAAATCGCATGCTGATCGCTTTGGCAAGAGGGTGGCAGAGCGCAAAAATCAAGA 750
Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAATCAAAAGTTCATCGGATCTACTGTTG 780
Db 201 LysSerLysValAspAlaIleThrTrpLeu 210

RESULT 9
Q8RJ83
ID Q8RJ83 PRELIMINARY; PRT; 210 AA.
AC Q8RJ83
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE RDXA protein.
GN RDXA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=345A; and 345C;
RA Bereswill S., Krainick C., Herrmann L., Kist M.;
RT "Absence of rdxA mutations in High-level Metronidazole Resistant
RL Clinical Isolates of Helicobacter pylori."
DR EMBL; AJ305352; CAC83804.1; -.
DR EMBL; AJ305353; CAC83805.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR02086; Aldehyde dehydr.
DR InterPro; IPR000415; Nitroreductase.
DR Pfam; PF00881; Nitroreductase; 1.
DR PROSITE; PS00070; ALDEHYDE DEHYDR. CYS; 1.
SQ SEQUENCE 210 AA; 24039 MW; 6F8104871026B981 CRC64;

Alignment Scores:
Pred. No.: 1.42e-99 Length: 210
Score: 1074.00 Matches: 207
Percent Similarity: 99.05% Conservative: 1
Best Local Similarity: 98.57% Mismatches: 2
Query Match: 68.89% Indels: 0
DB: 2 Gaps: 0

US-09-647-661-1 (1-900) x Q8RJ83 (1-210)

QY 151 ATGAAATTTTGGATCAGGAAAAAGAACAACTATTAAACGAGCGCATCTTTCGAAG 210
Db 1 MetLysPheLeuAspGlnGluLysArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTGACAGCCATTAGAGTTTCTAGTGAAGAAATAGAAATCGCTGAATCGCC 270

```

Db 21 MetPheAspSerHisTyrGluPheSerSerThrGluLeuGluGluLeuAlaGluIleAla 40
 QY 271 AGGCTATCGCCAGCTCTTACACACGACGCCATGGCATTTGTGTAGTGTACTATAAG 330
 Db 41 ArgLeuSerProSerTyrAsnThrGlnProThrPheValMetValThrAsnLys 60
 QY 331 GATTAAAAAACAATATGCAAGTCACAGCTACTTTAATGAAGAAATGATTAAGCGCT 390
 Db 61 AspLeuLysGlnIleAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
 QY 391 TCAGCGTAAATGGTGTATGCTCTTAAAGACCTAGCAGAGTGTACACACGCCATTAC 450
 Db 81 SerAlaLeuMetValValCysSerLeuLysProSerGluLeuLeuProHisGlyHisTyr 100
 QY 451 ATGCAAACTTACCGGAGTCTTATAAGTTAGAGTATCCCTCTTTTGTCTCAAAG 510
 Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
 QY 511 CTGGCGTGAAGTTCACACACGATGCAAGATTAAGAGCTATATTTAGAGCAATGC 570
 Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGluGlnCys 140
 QY 571 TATATCGCTGGGGCAATTCGATGGCGGTGAGCTTAATGGAGTGTAGTGTGCATT 630
 Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
 QY 631 ATTGGAGCTTGTATCCCTTTAAAGTGGTGAAGTTTGAAGAGCGTATCAATAAGCCT 690
 Db 161 IleGlyPheAspProLeuLysValGlyValLeuGluGluArgIleAsnLysPro 180
 QY 691 AAATCGCATGCTGTATCGCTTTGGGCAAGAGGTGCGCAGACGCGCAAAATCAAGA 750
 Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
 QY 751 AAATCAAAAGTGTATCGCTTACTTGGTTG 780
 Db 201 LysSerLysValAspAlaIleThrTripleu 210

RESULT 10

Q8RJCA ID Q8RJCA PRELIMINARY; PRT; 210 AA.
 AC Q8RJCA;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE RdxA protein.
 GN RDXA.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=210;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=053A, and 053C;
 RA Bereswill S., Krainick C., Herrmann L., Kist M.;
 RT "Absence of rdxA Mutations in High-Level Metronidazole Resistant
 RT Clinical Isolates of Helicobacter pylori."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ305348; CA83800.1;
 DR EMBL; AJ305349; CA83801.1;
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR002086; Aldehyde dehydr.
 DR InterPro; IPR000415; Nitroreductase.
 DR Pfam; PF00881; Nitroreductase; 1
 DR PROSITE; PS00070; ALDEHYDE_DEHYD CYS; 1.
 SQ SEQUENCE 210 AA; 24151 MW; 4BD0113D746F131A CRC64;

Alignment Scores:

Pred. No.: 1.79e-99 Length: 210
 Score: 1073.00 Matches: 207
 Percent Similarity: 99.05% Conservative: 1
 Best Local Similarity: 98.57% Mismatches: 2

Query Match: 68.83% Indels: 0
 DB: 2 Gaps: 0
 US-09-647-661-1 (1-900) x Q8RJCA (1-210)
 QY 151 ATGAAATTTTGGATCAGGAAAAAAGAACAACTATTAAACGAGCGCATTTCTTGCAAG 210
 Db 1 MetLysPheLeuAspGlnGluLysArgGlnLeuLeuAsnGluArgHisSerCysLys 20
 QY 211 ATGTTTGACACCCATTATGAGTTTCTAGTGAAGAATTAGAGAAATCGCTGAAATCGCC 270
 Db 21 MetPheAspSerHisTyrGluPheSerSerGluLeuGluGluIleAlaGluIleAla 40
 QY 271 AGGCTATCGCCAAAGCTTACACACGACGATCGCATTTTGTGTAGTGTGTACTATAAG 330
 Db 41 ArgLeuSerProSerSerTyrAsnThrGlnProThrPheValMetValThrAsnLys 60
 QY 331 GATTAAAAAACAATATGCAAGTTCACACACGATGCAAGATTAAGAGCTATATTTAGAGCAATGC 390
 Db 61 AspLeuLysGlnIleAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
 QY 391 TCAGCGTAAATGGTGTATGCTCTTAAAGACCTAGCAGAGTGTACACACGCCATTAC 450
 Db 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProTyrGlyHisTyr 100
 QY 451 ATGCAAACTTACCGGAGTCTTATAAGTTAGAGTATCCCTCTTTTGTCTCAAATG 510
 Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheThrGlnMet 120
 QY 511 CTGGCGTGAAGTTCACACACGATGCAAGATTAAGAGCTATATTTAGAGCAATGC 570
 Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGluGlnCys 140
 QY 571 TATATCGCTGGGGCAATTCGATGGCGGTGAGCTTAATGGAGTGTAGTGTGCATT 630
 Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
 QY 631 ATTGGAGCTTGTATCCCTTTAAAGTGGTGAAGTTTGAAGAGCGTATCAATAAGCCT 690
 Db 161 IleGlyGlyPheAspProLeuLysValGlyValLeuGluGluArgIleAsnLysPro 180
 QY 691 AAATCGCATGCTGTATCGCTTTGGGCAAGAGGTGCGCAGACGCGCAAAATCAAGA 750
 Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
 QY 751 AAATCAAAAGTGTATCGCTTACTTGGTTG 780
 Db 201 LysSerLysValAspAlaIleThrTripleu 210
 RESULT 11
 Q8FBC9 ID Q8FBC9 PRELIMINARY; PRT; 210 AA.
 AC Q8FBC9;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE RdxA.
 GN RDXA.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=210;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=951;
 RA MEDLINE=20356933; PubMed=10898705;
 RX Solca N.M., Bernasconi M.V., Fiffaretti J.C.;
 RT "Mechanism of metronidazole resistance in Helicobacter pylori:
 RT comparison of the rdxA gene sequences in 30 strains."
 RL Antimicrob. Agents Chemother. 44:2207-2210(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=951;

```

RA Maggi-Solca N., Piffaretti J.-C.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF180423; AAC01078.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002086; Aldehyde dehydr.
DR InterPro; IPR000415; Nitroreductase.
DR Pfam; PF00881; Nitroreductase; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
SQ SEQUENCE 210 AA; 24163 MW; 46C79907237A0207 CRC64;

Alignment Scores:
Pred. No.: 1.79e-99 Length: 210
Score: 1073.00 Matches: 207
Percent Similarity: 99.05% Conservative: 1
Best Local Similarity: 98.57% Mismatches: 2
Query Match: 68.83% Indels: 0
DB: 2 Gaps: 0

US-09-647-661-1 (1-900) x Q9FBC9 (1-210)
QY 151 ATGAAATTTTGGATCAGGAAAAAGAGACAACTATTAAACGAGCGCATCTTTCGAAG 210
Db 1 MetLysPheLeuAspGlnGluLysArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTTCACAGCCATTATGATTTCTAGTCAAGAAATTAGAGAAATCGCTGAATCGCC 270
Db 21 MetPheAspSerHisTyrGluPheSerSerGluGluLeuGluLeuAlaGluLeuAla 40
QY 271 AGCTATCGCCAGCTCTTCAACACGAGCCATGCGATTTGTGATGCTTACTATAAG 330
Db 41 ArgLeuSerProSerTyrAsnThrGlnProTrpHisPheValIleValThrAsnLys 60
QY 331 GATTTAAAAACAAATTCGAGTCCACAGCTACTTTAATGAAGAAATGATTAAGCGCT 390
Db 61 AspLeuLysLysGlnIleAlaValHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 391 TCAGCGTTAATGTTGATGCTTTTAAAGACCTTAGCGAGTGTGTACACGCGCCATTAC 450
Db 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProHisSerHisTyr 100
QY 451 ATGCAAAACCTTACCAGGAGTCTTAAAGTTAGAGTATGCTTCTTTTGTCAAATG 510
Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
QY 511 CTTCGCGTGAGATTCAACACGAGCATGCAAGATTAGAAAGCTATATTTTAGAGCAATGC 570
Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGluGlnCys 140
QY 571 TATATCGCTGTGGGCAATTTGCATGGCGGTAGCTTAATGGGATTGGATGCTTGCATT 630
Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGAGGCTTTGATCCTTTAAAAAGTGGGTGAAGTTTTAGAGAGCGGTATCAATAAGCCT 690
Db 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluArgIleAsnLysPro 180
QY 691 AAAATCGCATCTTGATCGCTTTGGCGAAGGGTGGCAGAGCGGCAAGCAAAAATCAAGA 750
Db 181 LysIleValCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAATCAAAAGTTGATGCGATTACTTGGTTG 780
Db 201 LysSerLysValAspAlaIleThrTripleu 210

RESULT 12
Q9VTP5 PRELIMINARY; PRT; 210 AA.
ID Q9VTP5
AC Q9VTP5;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Oxygen-insensitive NADPH nitroreductase.

```

```

GN RDXA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=220;
RA Marais A., Bilardi C., Cantet F., Megraud F.;
RT "Characterization of genes involved in the metronidazole resistance in
RT nine strains of Helicobacter pylori isolated before and after
RT eradication treatment.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF323006; AAL37279.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002086; Aldehyde dehydr.
DR InterPro; IPR000415; Nitroreductase.
DR Pfam; PF00881; Nitroreductase; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
SQ SEQUENCE 210 AA; 24035 MW; D22D60620BC2657 CRC64;

Alignment Scores:
Pred. No.: 1.79e-99 Length: 210
Score: 1073.00 Matches: 206
Percent Similarity: 99.52% Conservative: 3
Best Local Similarity: 98.10% Mismatches: 1
Query Match: 68.83% Indels: 0
DB: 2 Gaps: 0

US-09-647-661-1 (1-900) x Q8VTP5 (1-210)
QY 151 ATGAAATTTTGGATCAGGAAAAAGAGACAACTATTAAACGAGCGCATCTTTCGAAG 210
Db 1 MetLysPheLeuAspGlnGluLysArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTTCACAGCCATTATGATTTCTAGTCAAGAAATTAGAGAAATCGCTGAATCGCC 270
Db 21 MetPheAspSerHisTyrGluPheSerSerGluGluLeuGluLeuAlaGluLeuAla 40
QY 271 AGCTATCGCCAGCTCTTCAACACGAGCCATGCGATTTGTGATGCTTACTATAAG 330
Db 41 ArgLeuSerProSerTyrAsnThrGlnProTrpHisPheValIleValThrAsnLys 60
QY 331 GATTTAAAAACAAATTCGAGTCCACAGCTACTTTAATGAAGAAATGATTAAGCGCT 390
Db 61 AspValLysLysGlnIleAlaValHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 391 TCAGCGTTAATGCTGATGCTTTTAAAGACCTTAGCGAGTGTGTACACGCGCCATTAC 450
Db 81 SerAlaLeuMetValValCysSerLeuLysProSerGluLeuLeuProHisGlyHisTyr 100
QY 451 ATGCAAAACCTTACCAGGAGTCTTAAAGTTAGAGTATGCTTCTTTTGTCAAATG 510
Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
QY 511 CTTCGCGTGAGATTCAACACGAGCATGCAAGATTAGAAAGCTATATTTTAGAGCAATGC 570
Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGluGlnCys 140
QY 571 TATATCGCTGTGGGCAATTTGCATGGCGGTAGCTTAATGGGATTGGATGCTTGCATT 630
Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGAGGCTTTGATCCTTTAAAAAGTGGGTGAAGTTTTAGAGAGCGGTATCAATAAGCCT 690
Db 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluArgIleAsnLysPro 180
QY 691 AAAATCGCATCTTGATCGCTTTGGCGAAGGGTGGCAGAGCGGCAAGCAAAAATCAAGA 750
Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAATCAAAAGTTGATGCGATTACTTGGTTG 780

```



```

Db      201 LysSerLysValAspAlaIleThrTrpLeu 210
|||||
RESULT 13
Q8VTN9
ID      Q8VTN9          PRELIMINARY;      PRT;      210 AA.
AC      Q8VTN9;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Oxygen-insensitive NADPH nitroreductase.
GN      RDXA.
OS      Helicobacter pylori (Campylobacter pylori).
OC      Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC      Helicobacteraceae; Helicobacter.
OX      NCBI_TaxID=210;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CAS;
RA      Marais A., Bilardi C., Cantet F., Megraud F.;
RT      "Characterization of genes involved in the metronidazole resistance in
RT      nine strains of Helicobacter pylori isolated before and after
RT      eradication treatment.";
RL      Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF232014; AAL37287.1; -.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR002086; Aldehyde dehydr.
DR      InterPro; IPR000415; Nitroreductase.
DR      Pfam; PF00881; Nitroreductase; 1.
DR      PROSITE; PS02070; ALDEHYDE DEHYDR. CYS; 1.
SQ      SEQUENCE 210 AA; 24097 MW; 0ADC6F70463B4D0 CRC64;

Alignment Scores:
Pred. No.:      1.79e-99      Length:      210
Score:          1073.00      Matches:      207
Percent Similarity: 98.57%      Conservative: 0
Best Local Similarity: 98.57%      Mismatches: 3
Query Match:    68.93%      Indels:      0
DE:             2           Gaps:          0

US-09-647-661-1 (1-900) x Q8VTN9 (1-210)

QY      151 ATGAAATTTTGGATCAGGAAAAAGAACACACTATTAAACGAGCCCATTTCTTGAAG 210
Db      1 MetLysPheLeuAspGlnGluYsArgGlnLeuLeuAsnGluArgHisSerCysLys 20
|||||
QY      211 ATGTTTGACGCCATTATGAGTTTCTAGTCAAGAATTAGAAGAAATCGCTGAATCGCC 270
Db      21 MetPheAspSerHisTyrGluPheSerSerThrGluLeuGluGluIleAla 40
|||||
QY      271 AGGCTATCGCGAAGCTCTTACACACGCGCATGTCGATTTGTGATGCTTACTATAAG 330
Db      41 ArgLeuSerProSerSerTyrAsnThrGlnProThrHisPheValMetValThrAsnLys 60
|||||
QY      331 GATTTAAAAAACAATTCAGTCCAGCTACCTTAAATGAAGAAATGATTAAGCGCT 390
Db      61 AspLeuLysGlnIleAlaAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
|||||
QY      391 TCACGGTAAATGGTATGCTCTTTAAGACCTAGCGATGTTTACACAGCGCCATTAC 450
Db      81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTyr 100
|||||
QY      451 ATGCAGAACCTTACCCGGAGCTCTATAAGGTAGATGATCCCTTCTTTGCTCAATG 510
Db      101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheThrGlnMet 120
|||||
QY      511 CTGGCGTGAATTCACACAGCATGCAAGATAGAAAGCTATATTTTAGAGCAATGC 570
Db      121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGluGlnCys 140
|||||
QY      571 TATATCGCTGTGGGCAAAATTTGATGGCGGTGAGCTTAATGCGGATTCGATTCGATT 630
|||||

Db      141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
631 ATTCGAGGCTTTGATCCTTTAAAGTGGTGAAGCTTTTAGAGGCGGTATCAATAAGCCT 690
Db      161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluGluArgIleAsnLysPro 180
691 AAAATCGCATGCTTTCATCGCTTTGGCGAAGAGGGTGGCAGAGCGAGCCAAAAATCAAGA 750
Db      181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
751 AAATCAAAAGTTGATCGCATTCATTCTGTTG 780
Db      201 LysSerLysValAspAlaIleThrTrpLeu 210

RESULT 14
Q9FBD3
ID      Q9FBD3          PRELIMINARY;      PRT;      210 AA.
AC      Q9FBD3;
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      RDXA.
GN      RDXA.
OS      Helicobacter pylori (Campylobacter pylori).
OC      Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC      Helicobacteraceae; Helicobacter.
OX      NCBI_TaxID=210;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=56A;
RA      Solca N.M., Bernasconi M.V., Piffaretti J.C.;
RT      "Mechanism of metronidazole resistance in Helicobacter pylori:
RT      comparison of the rdxA gene sequences in 30 strains.";
RL      Antimicrob. Agents Chemother. 44:2207-2210(2000).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=56A;
RA      Maggi-Solca N., Piffaretti J.-C.;
RL      Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF180417; AAG01074.1; -.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR002086; Aldehyde dehydr.
DR      InterPro; IPR000415; Nitroreductase.
DR      Pfam; PF00881; Nitroreductase; 1.
DR      PROSITE; PS02070; ALDEHYDE DEHYDR. CYS; 1.
SQ      SEQUENCE 210 AA; 24110 MW; C77FBEE7B2C36F64 CRC64;

Alignment Scores:
Pred. No.:      2.25e-99      Length:      210
Score:          1072.00      Matches:      205
Percent Similarity: 99.05%      Conservative: 3
Best Local Similarity: 97.62%      Mismatches: 2
Query Match:    68.76%      Indels:      0
DE:             2           Gaps:          0

US-09-647-661-1 (1-900) x Q9FBD3 (1-210)

QY      151 ATGAAATTTTGGATCAGGAAAAAGAACACACTATTAAACGAGCCCATTTCTTGAAG 210
Db      1 MetLysPheLeuAspGlnGluYsArgGlnLeuLeuAsnGluArgHisSerCysLys 20
|||||
QY      211 ATGTTTGACGCCATTATGAGTTTCTAGTCAAGAATTAGAAGAAATCGCTGAATCGCC 270
Db      21 MetPheAspSerHisTyrGluPheSerSerThrGluLeuGluGluIleAla 40
|||||
QY      271 AGGCTATCGCGAAGCTCTTACACACGCGCATGTCGATTTGTGATGCTTACTATAAG 330
Db      41 ArgLeuSerProSerSerTyrAsnThrGlnProThrHisPheValMetValThrAsnLys 60
|||||
QY      331 GATTTAAAAAACAATTCAGTCCAGCTACCTTAAATGAAGAAATGATTAAGCGCT 390
|||||

```


Db 61 AspLeuLysGlnIleAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 391 TCAGCGTAATGGGTATGCTCTTAAGACCTTAGCGAGTTGTACACACGCCCAATTAC 450
Db 81 SerAlaLeuMetValCysSerLeuLysProSerGluLeuLeuProHisGlyHisTyr 100
QY 451 ATGCAAAACCTTTACCGGAGTCTTAAGCTTAGAGTGATCCCTCTTTTGCTCAAAATG 510
Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheThrGlnMet 120
QY 511 CTTCGCGTGAGATTCAACACAGCATGCAAGATAGAAAGCTATATTTTAGAGCAATGC 570
Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGluGlnCys 140
QY 571 TATATCGTGTGGGCGAAATTTGATGCGGTGAGCTTAATGGGATTGGATAGTTGCATT 630
Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGGAGCTTTGATCCCTTTAAAGTGGGTGAAGTTTATAGAGAGCTATCAATAAGCCT 690
Db 161 IleGlyGlyPheAspProLeuLysValGlyGluIleLeuGluGlnArgIleAsnLysPro 180
QY 691 AAAATCCGATCTGATCGCTTTGGCAAGAGGGTGGCAGAGCGGCCCAAAATCAAGA 750
Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAAATCAAAAGTTGATGCGATTACTTGTTG 780
Db 201 LysSerLysValAspAlaIleThrTrpLeu 210

RESULT 15
O25608 PRELIMINARY; PRT; 210 AA.
ID O25608
AC O25608;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Oxygen-insensitive NAD(P)H nitroreductase.
GN HP0354.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fieischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
DR EMBL; AE00604; RAD07997.1; --
DR FIR; B64639; B64639.
DR TIGR; HP0954; --
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002086; Aldehyde dehydr.
DR InterPro; IPR000415; Nitroreductase.
DR Efam; PF00881; Nitroreductase; 1
DR PROSITE; PS00070; ALDEHYDE DEHYDR_CYS; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 210 AA; 24068 MW; DFF15D2D102C002C CRC64;

Alignment Scores:
Pred. No.: 2,25e-99 Length: 210

Score: 1072.00 Matches: 207
Percent Similarity: 99.05% Conservative: 1
Best Local Similarity: 98.57% Mismatches: 2
Query Match: 68.76% Indels: 0
DB: 16 Gaps: 0

US-09-647-661-1 (1-900) x O25608 (1-210)

QY 151 ATGAATTTTGGATCAGGAAAAAAGACAACTATTAAACGAGCGCATTTCTTGAAG 210
Db 1 MetLysPheLeuAspGlnGluLysArgArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTTGACAGCCATTATGAGTTTCTAGTGAAGAAATAGAGAAATCCCTGAATCCGC 270
Db 21 MetPheAspSerHisTyrGluPheSerSerThrGluLeuGluIleAlaGluIleAla 40
QY 271 AGCTATCGCCCAAGCTCTTACAACACGAGCATGCGCATTTTGTGATCGTTACTAATAAG 330
Db 41 ArgLeuSerProSerSerTyrAsnThrGlnProTrpHisPheValMetValThrAspLys 60
QY 331 GATTTAAAAAACAATTCGAGTGCACAGCTACTTTAATGAAGAAATGATTAAGCGCT 390
Db 61 AspLeuLysLysGlnIleAlaAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 391 TCAGCGTTAATGCTGATGCTCTTTAAGACCTAGCGAGTTGTACACACGCCCAATTAC 450
Db 81 SerAlaLeuMetValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTyr 100
QY 451 ATGCAAAACCTTTACCGGAGTCTTTAAGGTTAGAGTGATCCCTCTTTTGCTCAAAATG 510
Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
QY 511 CTTCGCGTGAGATTCAACACAGCATGCAAGATAGAAAGCTATATTTTAGAGCAATGC 570
Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGluGlnCys 140
QY 571 TATATCGTGTGGGCGAAATTTGCGATGGCGGTGAGCTTAATGGGATTGGATAGTTGCATT 630
Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGGAGCTTTGATCCCTTTAAAGTGGGTGAAGTTTATAGAGAGCTATCAATAAGCCT 690
Db 161 IleGlyGlyPheAspProLeuLysValGlyGluIleLeuGluGlnArgIleAsnLysPro 180
QY 691 AAAATCGCATGCTTGATCGCTTTGGCAAGAGGGTGGCAGAGCGGCCCAAAATCAAGA 750
Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAAATCAAAAGTTGATGCGATTACTTGTTG 780
Db 201 LysSerLysValAspAlaIleThrTrpLeu 210

Search completed: April 20, 2004, 19:28:00
Job time : 72.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 20, 2004, 19:20:42 ; Search time 24.5 Seconds
(without alignments)
7067.133 Million cell updates/sec

Title: US-09-647-661-1
Perfect score: 1559
Sequence: 1 tgcagattttacagagacg.....gggggtttttcaagcgtttc 900

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Pgapop 6.0, Pgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xl2p
-C=/cgn2/1/USPTO_spool_p/US09647661/runat_20042004_135348_2706/app_query.fasta_1.1095
-DB=PIR_78 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -WATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09647661.ecgn 1.1.38 @runat_20042004_135348_2706 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREAGS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1072	68.8	210	2 B64639	oxygen-insensitive
2	1041	66.8	210	2 A71876	probable aldehyde
3	246	15.8	284	1 C64639	prolipo protein dia
4	229	14.7	283	2 B71876	probable prolipo
5	226	14.5	223	2 D90056	hypothetical prote
6	223.5	14.3	221	2 B69809	NAD(P)H-flavin oxi
7	223.5	14.3	221	2 B61149	probable NAD(P)H-f
8	218.5	14.0	220	1 B64114	probable NAD(P)H-f
9	208.5	13.4	221	2 C81155	NAD(P)H nitroreduc
10	203	13.0	222	2 E83700	NAD(P)H-flavin oxi
11	199.5	12.8	217	2 C82436	oxygen-insensitive
12	198.5	12.7	201	2 E81309	nitroreductase Cj1
13	198.5	12.7	217	1 A38686	nitroreductase (EC
14	196.5	12.6	217	1 S08397	nitroreductase (EC

ALIGNMENTS

RESULT 1

B64639

oxygen-insensitive NAD(P)H nitroreductase - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 11-Jun-1999
C;Accession: B64639

R;Comb, J.P.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujiki, C.; Bowman, C.; Watthey, J. Nature 388, 539-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, A.;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A;reference number: A64520, MUID:97394467; PMID:9252185

A;Accession: B64639

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-210 <TOM>

A;Cross-references: GB:AE000604; GB:AE000511; NID:g2314086; PIDN:AAD07997.1; PID:g2314

C;Superfamily: nitroreductase

Alignment Scores:

Pred. No.:	2.3e-93	Length:	210
Score:	1072.00	Matches:	207
Percent Similarity:	99.05%	Conservative:	1
Best Local Similarity:	98.57%	Mismatches:	2
Query Match:	68.76%	Indels:	0
DB:	2	Gaps:	0

US-09-647-661-1 (1-900) x B64639 (1-210)

Qy 151 ATGAATTTTGGATCAGGAAAGAGACAACTATTAAAGAGCGCCATTCTTCAAG 210

1 MetLysPheLeuAspGlnGluLysAArgGlnLeuLeuAsnGluAArgHisSerCysLys 20

Qy 211 ATCTTTGACGACCATATGAGTTTCTAGTGAAGATTAGAGAAATCGCTGAATCGGCC 270

21 MetPheAspSerHisTyrGluPheSerThrGluLeuGluCysLeuAlaGluIleAla 40

15 194.5 12.5 217 2 AD0573
16 191.5 12.3 217 2 H90705
17 191.5 12.3 217 2 C85556
18 190 12.2 217 2 H71912
19 188.5 12.1 217 1 I67685
20 185.5 11.9 202 2 H69902
21 182 11.7 217 1 B64600
22 181.5 11.6 201 2 S75047
23 172.5 11.1 209 2 C69783
24 169.5 10.9 321 2 B72384
25 167.5 10.7 271 2 B98179
26 167 10.7 213 2 B87429
27 165.5 10.6 213 2 D83929
28 164.5 10.6 209 2 B93108
29 160 10.3 201 2 B95072
30 159 10.2 210 2 A81087
31 157.5 10.1 201 2 B97940
32 153.5 9.8 179 2 G97074
33 153.5 9.8 356 2 H86808
34 153 9.8 174 1 C69533
35 153 9.8 215 2 A75331
36 145.5 9.3 212 2 C84132
37 144.5 9.3 203 2 D86647
38 144 9.2 210 2 A81451
39 143 9.2 213 2 A84250
40 140 9.0 208 2 G97004
41 139.5 8.9 200 2 E82996
42 137.5 8.8 206 2 D69781
43 129.5 8.3 194 2 C89266
44 126 8.1 191 1 B99027
45 126 8.1 249 2 AB2039

oxygen-insensitive
oxygen-insensitive
oxygen-insensitive
probable oxidoredu
nitroreductase (EC
nitroreductase hom
NAD(P)H-flavin oxi
drgA protein - Syn
NADH dehydrogenase
bacterioferritin c
probable NAD(P)H n
NAD(P)H-flavin oxi
NAD(P)H nitroreduc
nitroreductase fam
nitroreductase (im
nitroreductase fam
NADH oxidase homol
nitroreductase (EC
nitroreductase fam
transcription regu
NAD(P)H-flavin oxi
nitroreductase - D
NADH dehydrogenase
oxidoreductase ybi
NADH oxidase homol
NADH oxidase (limpo
nitroreductase fam
probable nitroredu
conserved hypothet
NAD(P)H-flavin oxi
NADPH-oxidoreducta
hypothetical prote

QY 271 AGGCTATCGCAAGCTCTTAAACACGACGACCGATTTTGATGGTTACTAATAAG 330
 Db |||||
 QY 41 ArgLeuSerProSerSerTyAsnThrGlnProTriPhisPheValMetValThrAspLys 60
 Db |||||
 QY 331 GATTATAAAAAAATAATGCGTCCACAGCTACTTTAATGAAGAATGATTAAGCGCT 390
 Db |||||
 QY 61 AspLeuLysGlnIleAlaAlaHisSerTyPheAsnGluGluMetIleLysSerAla 80
 Db |||||
 QY 391 TCAGCGTTAATGGTGGTATGCTCTTAAAGACCTAGCGAGTTGTTACACACGCGCCATTAC 450
 Db |||||
 QY 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTy 100
 Db |||||
 QY 451 ATGCAAAACCTTTACCGGAGCTTTAAGTTAGATGATGATCCCTCTTTGCTCAAAATG 510
 Db |||||
 QY 101 MetGlnAsnLeuTyProGluSerTyLysValArgValIleProSerPheAlaGlnMet 120
 Db |||||
 QY 511 CTTCGGCTGAGATTAACACAGCATGCAAGATAGAAAGCTATATTTTAGAGCAATGC 570
 Db |||||
 QY 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyLleLeuGluGlnCys 140
 Db |||||
 QY 571 TATATCGCTGGGGCAAAATTTGATGGGCTGAGCTTAATGGGATTTGATGATTCGATT 630
 Db |||||
 QY 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
 Db |||||
 QY 631 ATTGGAGCTTTGATCCTTTAAAGTGGTGAGCTTTTAGAGAGCGTATCAATAGCCT 690
 Db |||||
 QY 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluArgIleAsnLysPro 180
 Db |||||
 QY 691 AAAATCGCATCTTGATCGCTTTGGCAAGAGGTGGCAGAGCGCAAGCAAAATCAAGA 750
 Db |||||
 QY 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
 Db |||||
 QY 751 AAATCAAACTTGATCGCATCTTGCTGTTG 780
 Db |||||
 QY 201 LysSerLysValAspAlaIleThrTrpLeu 210
 Db |||||

RESULT 2
 A:1876
 Probable aldehyde dehydrogenase - Helicobacter pylori (strain J99)
 C:Species: Helicobacter pylori
 A:Variety: strain J99
 C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 11-Jun-1999
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
 Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
 Nature 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
 A:Reference number: A71800; PMID:99120557; PMID:9923682
 A:Accession: A71876
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-210 <ARN>
 A:Cross-references: GB:AE001518; GB:AE001439; NID:g4155454; PIDN:ABD06472.1; PID:g415546
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: jnp0888
 C:Superfamily: nitroreductase

Alignment Scores:
 Pred. No.: 1,98e-90 Length: 210
 Score: 1041.00 Matches: 200
 Percent Similarity: 97.14% Conservative: 4
 Best Local Similarity: 95.24% Mismatches: 6
 Query Match: 66.77% Indels: 0
 DB: 2 Gaps: 0

US-09-647-661-1 (1-900) x A71876 (1-210)

QY 151 ATGAATTTTGGATCAGAAAAGACAGCACTATTAAACGCGCCATTCTTGAAG 210
 Db |||||
 QY 1 MetLysPheLeuAspHisGlnLysArgGlnLeuLeuAsnGluArgHisSerCysLys 20
 Db |||||
 QY 211 ATGTTTGACGCCATTATGAGTTTCTAGTGAAGAATTAGAAGAAATCGCTGAAATCGCC 270
 Db |||||

Db 21 MetPheAspSerHisTyGluPheSerSerThrGluLeuGluGluIleAlaGluIleAla 40
 QY |||||
 QY 271 AGGCTATCGCAAGCTCTTAAACACGACGACCGATTTTGATGGTTACTAATAAG 330
 Db |||||
 QY 41 ArgLeuSerProSerSerTyAsnThrGlnProTriPhisPheValMetValThrAsnLys 60
 Db |||||
 QY 331 GATTATAAAAAAATAATGCGTCCACAGCTACTTTAATGAAGAATGATTAAGCGCT 390
 Db |||||
 QY 61 AspLeuLysGlnIleAlaAlaHisSerTyPheAsnGluGluMetIleLysSerAla 80
 Db |||||
 QY 391 TCAGCGTTAATGGTGGTATGCTCTTAAAGACCTAGCGAGTTGTTACACACGCGCCATTAC 450
 Db |||||
 QY 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTy 100
 Db |||||
 QY 451 ATGCAAAACCTTTACCGGAGCTTTAAGTTAGATGATGATCCCTCTTTGCTCAAAATG 510
 Db |||||
 QY 101 MetGlnAsnLeuTyProGluSerTyLysValArgValIleProSerPheAlaGlnMet 120
 Db |||||
 QY 511 CTTCGGCTGAGATTAACACAGCATGCAAGATAGAAAGCTATATTTTAGAGCAATGC 570
 Db |||||
 QY 121 LeuGlyValArgPheAsnHisSerMetGlnLysLeuGluSerTyLleLeuGluGlnCys 140
 Db |||||
 QY 571 TATATCGCTGGGGCAAAATTTGATGGGCTGAGCTTAATGGGATTTGATGATTCGATT 630
 Db |||||
 QY 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
 Db |||||
 QY 631 ATTGGAGCTTTGATCCTTTAAAGTGGTGAGCTTTTAGAGAGCGTATCAATAGCCT 690
 Db |||||
 QY 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluArgIleAsnLysPro 180
 Db |||||
 QY 691 AAAATCGCATCTTGATCGCTTTGGCAAGAGGTGGCAGAGCGCAAGCAAAATCAAGA 750
 Db |||||
 QY 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerLysLysSerArg 200
 Db |||||
 QY 751 AAATCAAACTTGATCGCATCTTGCTGTTG 780
 Db |||||
 QY 201 LysSerLysValAspAlaIleThrTrpLeu 210
 Db |||||

RESULT 3
 C64639
 prolipoprotein diacylglycerol transferase (EC 2.4.99.-) - Helicobacter pylori (strain :
 C:Species: Helicobacter pylori
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.I.
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, I.
 Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; PMID:97394467; PMID:9252185
 A:Accession: C64639
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-284 <TOM>
 A:Cross-references: GB:AE000604; GB:AE000511; NID:g2314086; PIDN:ABD07998.1; PID:g2314
 C:Superfamily: prolipoprotein diacylglycerol transferase
 C:Keywords: glycosyltransferase

Alignment Scores:
 Pred. No.: 3.7e-15 Length: 284
 Score: 246.00 Matches: 49
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.00% Mismatches: 0
 Query Match: 15.78% Indels: 0
 DB: 1 Gaps: 0

US-09-647-661-1 (1-900) x C64639 (1-284)

QY 2 CGAATTTTACAGAGCGCAGATAGCCAAATCGGGTTATTTTAAATTTGACATG 61
 Db |||||
 QY 235 AlaGluPheTyArgGluProAspSerGlnMetGlyValTyPheLeuAsnLysSerMet 254
 Db |||||

QY 62 GGCAGATTTTAACTTATTATGAGTGTGTTTCATTAGGAGTTTATTATGATGCTACA 121
Db 255 GlyGlnIleLeuSerLeuPheMetValIleValSerLeuGlyIleLeuLeuTyAlaThr 274
QY 122 AAAAATCTTAAATAAAGAAATCAAA 151
Db 275 LysAsnSerLysLysIleLysGluAsnGln 284

RESULT 4
B71876
probable prolipoprotein diacylglycerol transferase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 18-Jun-1999
C:Accession: B71876
; Alm, K.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923692
A:Accession: B71876
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-283 <ARN>
A:Cross-references: GB:AE001439; NID:G4155454; PIDN:AA06473.1; PID:G415547
A:Experimental source: strain J99
C:Genetics:
A:Gene: lgt
C:Superfamily: prolipoprotein diacylglycerol transferase

Alignment Scores:
Pred. No.: 1.5e-13 Length: 283
Score: 229.00 Matches: 45
Percent Similarity: 100.00% Conservatives: 3
Best Local Similarity: 93.75% Mismatches: 0
Query Match: 14.69% Indels: 0
DB: 2 Gaps: 0

US-09-647-661-1 (1-900) x B71876 (1-283)
QY 2 GCAGATTTTACAGAGCCAGATAGCCAAATGGGGTATTATTTTAAATTTGAGCATG 61
Db 235 AlaGluPheTyArgGluProAspSerGlnLeuGlyValTyPheLeuAsnLeuSerMet 254
QY 62 GGCAGATTTTAACTTATTATGAGTGTGTTTCATTAGGAGTTTATTATGATGCTACA 121
Db 255 GlyGlnIleLeuSerValPheMetValIleValSerLeuGlyIleLeuLeuTyAlaThr 274
QY 122 AAAAATCTTAAATAAAGAAATCAAA 145
Db 275 LysAsnSerLysLysIleLysGlu 282

RESULT 5
D90056
hypothetical protein SA2311 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 01-Mar-2002
C:Accession: D90056
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A99758; MUID:21311952; PMID:11418146
A:Accession: D90056
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-223 <KUR>
A:Cross-references: GB:BA000018; PID:g13702473; PIDN:BA043614.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA2311

C:Superfamily: nitroreductase
Alignment Scores:
Pred. No.: 2.84e-13 Length: 223
Score: 226.00 Matches: 55
Percent Similarity: 51.21% Conservatives: 51
Best Local Similarity: 26.57% Mismatches: 87
Query Match: 14.50% Indels: 14
DB: 2 Gaps: 5

US-09-647-661-1 (1-900) x D90056 (1-223)
QY 196 GCGCATTTCTGCAGATGTTTTCAGACGCGCATTTATGAGTTTCTAGTGAAGAAATTAGAAGAA 255
Db 15 ArgHisAlaThrLysGlnPheAspProGlnLysLysValSerLysGluAspPheGluThr 34
QY 256 ATCGCTGAATTCGCCAGCTATCGCAGCTCTTACACACGCGAGCCATGCGATTTCGTG 315
Db 35 IleLeuGluSerGlyArgLeuSerProSerSerLeuGlyLeuGluProTrpLysPheVal 54
QY 316 ATGGTTTACTAATAGGATTTAAAAACAAATTCGAGTCGACGACTTCTTAAATGAAGAA 375
Db 55 ValIleGlnAspGlnAlaLeuArgAspGluLeuLysAlaHisSerTrpGlyAlaAlaLys 74
QY 376 ATGATTAAGCGCTTCAGCTTAATGCTGTGTTATGCTCTTTAAGACCTAGGAGTTGTTA 435
Db 75 GlnLeuAspThrAlaSerHisPheValLeuIlePheAlaArgLys-----AsnValThr 92
QY 436 CCACACGCGCATTTACATGCAAAACCTTTACCGGAG-----TCTTATAAGGTTAGAGTG 489
Db 93 SerArgSerProTyValGlnHisMetLeuArgAspIleLysLysTyArgGluAlaGlnThr 112
QY 490 ATCCCTTCTTTTGTCTCAAAATGCTTGGCGTG-----AGATTCAACACACAGC 534
Db 113 IleProAlaValGluGlnLysPheAspAlaPheGlnAlaAspPheHisIleSerAspAsn 132
QY 535 ATGCAAGATTTAGAAAGCTATATTTCAGCAATGCTATATCGTGTGCGGCAATTTGC 594
Db 133 AspGlnAlaLeuTyArgTrpSerSerLysGlnThrTyIleAlaLeuGlyAsnMetMet 152
QY 595 ATGGCGGTGAGCTTAATCGGATTCGATAGTTGCATATTATGGAGGCTTTTCATCTTTAAAA 654
Db 153 ThrThrAlaAlaLeuLeuGlyIleAspSerCysProMetGluGlyPheSerLeuAspThr 172
QY 655 GTGGGTGAAGTTTGAAGAGCGT-----ATCAATAAGCCCTAAATCGCATGC 702
Db 173 ValThrAspIleLeuAlaAsnLysGlyIleLeuAspThrGluGlnPheGlyLeuSerVal 192
QY 703 TTGATCGCTTTGGCAAGAGGTGGCAGAA---GCGAGCCAAATCAAGAAATCAAAA 759
Db 193 MetValAlaPheGlyTyArgGlnGlnAspProProLysAsnLysThrArgGlnAlaTy 212
QY 760 GTTGATCGCATTTACTTGGTTG 780
Db 213 GluAspValIleGluTrpVal 219

RESULT 6
B69809
NAD(P)H-flavin oxidoreductase homolog yfko - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: B69809
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bert
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari, E
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y., M.; Ogasawa, K.; Ogawa, K.; Oudega, B.; Park, S.H.; Paro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlo
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.P.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: B69809
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-221 <KON>
A:Cross-references: GB:Z39108; GB:AL009136; NID:g2633055; PIDN:CAB12612.1; PID:g2633107
A:Experimental source: strain 168
C:Genetics:
A:Gene: yfkO
C:Superfamily: nitroreductase

Alignment Scores:
Pred. No.: 4,89e-13 Length: 221
Score: 223.50 Matches: 57
Percent Similarity: 49.09% Conservative: 51
Best Local Similarity: 25.91% Mismatches: 91
Query Match: 14.34% Indels: 21
DB: 2 Gaps: 4

US-09-647-661-1 (1-900) x B69809 (1-221)

QY 175 AGAAGACCACTATTA-----AACGAGCGCCATCTTCAAGATGTTTGACAGCCAT 225
DB 5 LysThrGlnIleLeuAspAlaTyrAsnPhaGHisAlaThrLysGluPheAspProAsn 24
QY 226 TATGAGTTTCTAGTGAAGATTAGAGAAATCGCTGAATCGCCAGGCTATCGCCAAAGC 285
DB 25 LysIleValSerAspSerAspGluPheIleLeuGluThrGlyArgLeuSerProSer 44
QY 286 TCTTACACACGCGCCATGCTTTGTGATGTTCTAATAGGATTAAAGAAACAA 345
DB 45 SerLeuGlyLeuGluProTyrPheValValGlnAsnProGluPheArgGluLys 64
QY 346 ATTGCAGTGCACAGCTACTTAAATCAAGAAATGATTAAGCGCTTCAGCGTTAATGGTG 405
DB 65 LeuArgGluTyrThrTrpGlyAlaGlnLysGlnLeuProThrAlaSerHisPheValLeu 84
QY 406 GTATGCTCTTAAGACCTAGCGAGTTGTATACACAGCCCATACATCGCAAAACCTTTAC 465
DB 85 IleLeuAlaArgThrAlaLysAspIleLysTyrAsnAlaAspTyrIleLysArgHisLeu 104
QY 466 CCGAGTCTTAAAGGTTAGATGATCCCTCTTTGCTCAATGCTTGGCGTGAGATTC 525
DB 105 LysGlu-----ValLysGlnMetProGlnAspValTyrGluGlyTyrLeuSerLys 121
QY 526 AACACACATGCAAGATTAGAAGCTATATTTAGAG----- 564
DB 122 ThrGluGlyPheGlnLysAsnAspLeuHisLeuGluSerAspArgThrLeuPheAsp 141
QY 565 -----CAATGCTATATCGCTGTGGGCAAAATTTGCATGGCGGTGAGCTTAATG 612
DB 142 TrpAlaSerLysGlnThrTyrIleAlaLeuGlyAsnMetMetThrAlaAlaGluIle 161
QY 613 CGATTGGATGATTGAGGCTTTCATCTTAAAGTGGGTGAAGTTTAAAG 672
DB 162 GlyValAspSerCysProIleGluGlyPheGlnTyrAspHisIleHisArgIleLeuGlu 181
QY 673 GAG-----CGTATCAATAAGCTAAATCGCATCTGATCGCTTTCGGCAAG 720
DB 182 GluGluGlyLeuLeuGluAsnGlySerPheAspIleSerValMetAlaAlaPheGlyTyr 201
QY 721 AGGTTGGCAGAGCGCAAAATCAAGAAATCAAGAAATGATGATGCTTCTGTTG 780
DB 202 ArgValArgAspProArgProLysThrArgSerAlaValGluAspValValLysTrpVal 221

RESULT 7
E81949
probable NAD(P)H-flavin oxidoreductase NMA1015 [imported] - Neisseria meningitidis (stra
C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: E81949
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; More
; Holroyd, S.; Jags, K.; Leather, S.; Mould, S.; Mungall, K.; Quail, M.A.; Rajandran
Nature 404, 502-505, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: A61775; MUID:20222556; PMID:10761919
A:Accession: E81949
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-221 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84284.1; PID:g73797
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1015
C:Superfamily: nitroreductase

Alignment Scores:
Pred. No.: 4,89e-13 Length: 221
Score: 223.50 Matches: 58
Percent Similarity: 47.79% Conservative: 50
Best Local Similarity: 25.66% Mismatches: 95
Query Match: 14.34% Indels: 23
DB: 2 Gaps: 4

US-09-647-661-1 (1-900) x E81949 (1-221)

QY 151 ATGAAATTTTGGATCAGGAAAGAGACAACTATTAAAGCGCCATCTTGGCAAG 210
DB 1 MetThrValLeuAspArgGlnGlnValLeuSerAlaPheLysAsnArgLysSerCysArg 20
QY 211 ATGTTTGACAGCCATATGAGTTTCTAGTGAAGAAATAGAGAAATCGCTGAATCGCC 270
DB 21 HisTyrAspAlaAlaArgLysIleSerAlaGluAspPheGlnPheIleLeuGluLeuGly 40
QY 271 AGCTATCGCCAAAGCTTTTACACAGCGCCATGCGATTTGCTGATGCTTACTTAAG 330
DB 41 ArgLeuSerProSerValGlySerGluProTyrGlnPheValValGlnAsnPro 60
QY 331 GATTTAAAGAAACAAATTCGCTGACAGCTACTTTAATGAAGAAATGATTAAGAAACGCT 390
DB 61 GluIleArgGlnAlaIleLysProPheSerTrpGlyMetAlaAspAlaLeuAspThrAla 80
QY 391 TCAGCGTTAATGCTGGTATGCTCTTTAAGACCTAGCGATTTGTTACACAGCGCCATTAC 450
DB 81 SerHisLeuValValPheLeuAlaLysLys-----AsnAlaArgPhe 94
QY 451 ATGCAAAACCTTTACCGGAGTCTTTAAGGTTAGAGTGATC-----CCTTCTTTT 501
DB 95 AspSerProPheMetLeuGluSerLeuLysArgAlaGlyValThrGluProAspAlaMet 114
QY 502 GTCATAATGCTTGGC-----GTGAGATTCAACCAACAGC 534
DB 115 AlaLysSerLeuAlaArgTyrGlnAlaPheGlnAlaAspAspIleLysIleLeuAspAsp 134
QY 535 ATGCAAGATTAGAAAGCTATATTTAGACCAATCTATCGCTGTGGGCAAAATTTGC 594
DB 135 SerArgAlaLeuPheAspTrpCysArgGlnThrTyrIleAlaLeuGlyAsnMetMet 154
QY 595 ATGGGCGTGAGCTTAATGGATGATGATGCTTATTTGGAGGCTTTCATCTTCTTAAAG 654
DB 155 ThrGlyAlaAlaMetAlaGlyIleAspSerCysProValGluGlyPheAsnTyrAlaAsp 174
QY 655 GTGGGTGAAGTTTGAAGAGCGTATCAAT-----AAGCCTAAATCGCA 699
DB 175 MetGluArgValLeuSerGlyGlnPheGlyLeuPheAspAlaAlaGluTrpGlyValSer 194
QY 700 TGGTTGATCGCTTGGCAAGAGGTTGGCAGAGCGAGCCAGAAATCAAGAAATCAAGAA 759
DB 195 ValAlaAlaThrPheGlyTyrArgValGlnGluIleAlaThrLysAlaArgProLeu 214
QY 760 GTTGATCGCATTAATTGG 777

Db 215 GluGluThrValIleTrp 220

RESULT 8
B64114
probable NAD(P)H-flavin oxidoreductase (EC 1.6.6.-) - Haemophilus influenzae (strain Rd
C;Species: Haemophilus influenzae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: B64114
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A.;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: B64114
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-220 <TR>
A;Cross-references: GB:U32807; GB:I42023; NID:91574725; PIDN:AAC22926.1; PID:g1574733; T
A;Note: named as homolog to a protein from Vibrio fischeri
C;Superfamily: nitroreductase
C;Keywords: NAD; oxidoreductase

Alignment Scores:
Pred. No.: 1.45e-12 Length: 220
Score: 218.50 Matches: 59
Percent Similarity: 48.89% Conservative: 51
Best Local Similarity: 26.22% Mismatches: 94
Query Match: 14.02% Indels: 21
DB: 1 Gaps: 4

US-09-647-661-1 (1-900) x B64114 (1-220)

QY 151 ATGAAATTTTGGATCAGGAAAAAAGACAACTATTAAACGAGCGCCATCTTTCGAAG 210
Db 1 MetThrGlnLeuThrArgGluGlnValLeuGluLeuPheHisGlnArgSerThrArg 20
QY 211 ATGTTTGACGACCATATGAGTTTCTAGTGAAGANTTAGAAGAAATCGCTGAATCGCC 270
Db 21 TyrTyrAspProThrLysLysIleSerAspGluAspPheGluCysIleLeuGluCysGly 40
QY 271 AGGCTATCGCAAGCTCTTACACAGCGACGCCATCGCATTTTGTGATGTTACTAATAAG 330
Db 41 ArgLeuSerProSerValGlySerGluProTyrIlePheLeuValIleGlnAsnLys 60
QY 331 GATTAAAAAACAATTCGACGACGACTATTAATGAAGAAATGATTAAGCGCT 390
Db 61 ThrLeuArgGluLysMetLysProPheSerTrpGlyMetIleAsnGlnLeuAspAsnCys 80
QY 391 TCAGCGTAAATGTTGGTATGCTCTTAAAGACCTAGCGAGTTGTACCACGCGCATAC 450
Db 81 SerHisLeuValIleLeuAlaLysLys-----AsnAlaArgTyr 94
QY 451 ATGCAAAACCTTACCGGAGTCTTAAAGTTAGATGATGATCCCTTCTTTTGTCAAATG 510
Db 95 AspSerProPheValAspValMetAlaArgLysGlyLeuAsnAlaGluGlnGlnGln 114
QY 511 CTGGCGTGAGATCAACACAGCATGCAAGATTAGAAGCTATATTTTAGAG----- 564
Db 115 AlaAlaLeuThrLysTyrLysAlaLeuGlnGluAspMetLysLeuLeuGluAsnAsp 134
QY 565 -----CAATGCTATATGCTGCTGGGCAAAATTTGCATG 597
Db 135 ArgThrLeuPheAspTrpCysSerLysGlnThrTyrIleAlaLeuAlaAsnMetLeuThr 154
QY 598 GCGGTGAGCTTAATGGATTGGATGATGTCATTTAGGAGGCTTTCATCCTTTAAAGTG 657
Db 155 GlyAlaSerAlaLeuGlyIleAspSerCysProIleGluGlyPheHisTyrAspLysMet 174
QY 658 GGTGAAGCTTTA--GAAGAGCGGTATCAATAAGCTAAA-----ATCGCATGCTTG 705
Db 175 AsnGluCysLeuAlaGluGluGlyLeuPheAspProGlnGluTyrAlaValSerValAla 194

QY 706 ATCGCTTTGGCAAGAGCGTGGCAGAGCGAGCGCAAAAATCAAGAAATCAAAAGTTGAT 765
Db 195 AlaThrPheGlyTyrArgSerArgAspIleAlaLysLysSerArgLysGlyLeuAspGlu 214
QY 766 GCGATTACTTGGTTG 780
Db 215 ValValLysTrpVal 219

RESULT 9
C81155
NAD(P)H nitroreductase, probable NMB0804 [imported] - Neisseria meningitidis (strain MC
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: C81155
R;Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizzia, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: C81155
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-221 <TET>
A;Cross-references: GB:AF002435; GB:AF002098; NID:g7226049; PIDN:AAF41217.1; PID:g72260
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0804
C;Superfamily: nitroreductase

Alignment Scores:
Pred. No.: 1.29e-11 Length: 221
Score: 208.50 Matches: 58
Percent Similarity: 46.46% Conservative: 47
Best Local Similarity: 25.66% Mismatches: 98
Query Match: 13.37% Indels: 23
DB: 2 Gaps: 5

US-09-647-661-1 (1-900) x C81155 (1-221)

QY 151 ATGAAATTTTGGATCAGGAAAAAAGACAACTATTAAACGAGCGCCATCTTTCGAAG 210
Db 1 MetThrValLeuSerLysGluGlnValLeuSerAlaPheLysAsnArgLysSerCysArg 20
QY 211 ATGTTTGACGCCATATGAGTTTCTAGTGAAGAAATGAGAAATCGCTGAATCGCC 270
Db 21 HisTyrAspAlaAlaArgLysIleSerAlaGluAspPheGlnPheIleLeuGluLeuGly 40
QY 271 AGGCTATCGCAAGCTCTTACACAGCGACGCCATCGCATTTTGTGATGTTACTAATAAG 330
Db 41 ArgLeuSerProSerValGlySerGluProTyrIlePheLeuValIleValGlnAsnPro 60
QY 331 GATTAAAAAACAATTCGACGACGACTATTAAGCTATGCTCTTAAAGACCTAGCGAGTTGTACCACGCGCATAC 450
Db 61 GluIleArgGlnAlaIleLysProPheSerTrpGlyMetAlaAspAlaLeuAspThrAla 80
QY 391 TCAGCGTAAATGTTGGTATGCTCTTAAAGACCTAGCGAGTTGTACCACGCGCATAC 450
Db 81 SerHisLeuValValPhe-----LeuAlaLysLysAsnAlaArgSerAspSerProPhe 98
QY 451 ATGCAAAACCTTACCGGAGTCTTATAAGGTTTGA-----GTG 489
Db 99 MetLeu-----GluSerLeuLysArgGlyValThrGluProAspAlaVal 114
QY 490 ATCCCTTCTTTTGGT-----CAATGCTGGCGTGAGATTCACACACAGC 534
Db 115 AlaLysSerLeuAlaArgTyrGlnAlaPheGlnAlaAspAspIleLysIleLeuAspAsp 134
QY 535 ATGCAAGATTAGAAAGCTATATTTAGACGAATGCTATATCGCTGTGGGCAAAATTCG 594
Db 135 SerArgAlaLeuPheAspTrpCysArgGlnThrTyrIleAlaLeuAlaAsnMetMet 154

QY 595 ATGGCGCTGAGCTTAATCGCATTCGATTGCATTATTTGAGCGCTTGTATCCTTTAAAAA 654
Db : : : : :
Dy ThrGlyAlaalaMetAlaGlyIleAspSerCysProValgluGlyPheAsnTyraLalGlu 174
QY 655 GTGGGTGAAGTTTTAGAAGACGGTATCAAT-----BAGCCATAAAATCGCA 699
Db : : : : :
Dy MetGluArgIleLeuSerGlyGlnPheGlyLeuPheAspAlaalaGluTrpGlyValSer 194
QY 700 TGGTTGATCGCTTTGGCAAGAGGGTGCGAGACGCGCAAAAATCAAGAAAATCAAAA 759
Db : : : : :
Dy ValAlaalaThrPheGlyTyArgValGlnGluileAlaThrLysAlaA-GArgProLeu 214
QY 760 GTTCATCGCATTTACTTCG 777
Db : : : : :
Dy GluGluThrValIleTrp 220

RESULT 10
E83700
NAD(P)H-flavin oxidoreductase BH0405 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Mar-2002
C;Accession: E83700
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai,
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83550; MUID:20512582; PMID:11058132
A;Accession: E83700
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <STO>
A;Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:SAB04124.1; GSPDB:GNOC
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0405
C;Superfamily: nitroreductase

Alignment Scores:
Pred. No.: 4.27e-11 Length: 222
Score: 203.00 Matches: 56
Percent Similarity: 50.46% Conservative: 54
Best Local Similarity: 25.69% Mismatches: 88
Query Match: 13.02% Indels: 20
DB: 2 Gaps: 7

US-09-647-661-1 (1-900) x E83700 (1-222)

QY 175 AGAAGACAACCTATTAAACGAG-----CGCCATTCTTCCAGATGTTTGACAGCCAT 225
Db : : : : :
Dy LysGluGluileLeuAsnAlaTyArgPheArgHisAlaThrLysGluPheAspProasn 27
QY 226 TATGAGTTTTCTAGTGAAGATTAAGAAGAACCTCGTAAATCGCAGGCTATCGCCAAGC 285
Db : : : : :
Dy LysThrileProAlaAspAspGlnPheileLeuGluThrGlyArgLeuSerProser 47
QY 286 TCITTACACAGCGCCATGCCATTTCGTGNTGTTTACTTAATAAGGATTTAAAAACAA 345
Db : : : : :
Dy SerPheGlyMetGluProTrpArgPheileValGlnAsnGluLysLeuArgAspLys 67
QY 346 ATTGCAGTGCACGCTACTTTTAATGAAGAAATGATTAAGACGGCTTCAGCGTTTAAATGGTG 405
Db : : : : :
Dy IleLysasnThrSerTrpGlyAlaAlaGlyLysLeuProGluAlaSerHisPheValMet 87
QY 406 GTATGCTCTTTAAGACCTTAGCGATTTGTATCCACAGCGCCATTACATGCAAAACCTTTAC 465
Db : : : : :
Dy IleLeuAlaArgrThrHisLeuAspThrArgPheAspSerAlatyrLeuLysGluGlnPhe 107
QY 466 CCGAGTCTTATAGGTTAGAGTGCCTCTTTTCCTCAAATGCTTGGCGTG----- 519
Db : : : : :
Dy SerAspPheProGluAspArgLeu-----SerGlnTyLeuglyMetileclu 123
QY 520 AGATTC---AAACCACGATCGCAAGATTAGAA-----AGCTATATT 558
Db : : : : :
Dy : : : : :

Db 124 GlnPheGlnLysHisAspPheLeuLeuAspGlnGluArgPheLeuAsnAspTrpAla 143

Qy 559 TTAGAGCAATGCCTATATCGCTGTGGGCCAAATTTTCATGGCGTGACGTTAATGGATTG 618
:: ||||| ::
Db 144 IleArgGlnThrTyrlleProLeuAlaAsnMetMetThrAlaAlaGluIleGlyTlle 163
:: ||||| ::
Qy 619 GATAGTTGCATTATTCGAGGCTTTGATCCCTTTAAAGTGGGTGAAGTTTA--CAAGAG 675
||||| :::::: ::|||
Db 164 AspSerCysProIleGluGlyPheAsnValGluArgMetAsnGluLeuLeuAlaGluGlu 183
||||| :::::: ::|||
Qy 676 CGTATCAATAAGCCCT-----AAAATCGCATGCTTGATCGCTTTGGCGCAAGAGGTTG 726
::: ::: |||:::
Db 184 GlyLeuLeuGluAspGlyHisPheThrIleSerValMetValAlaPheGlyTyrArgLys 203
::: ::: |||:::
Qy 727 GCAGAAGCGAGCCAAAAATCAGAAAATCAAAGTTGATCGATACTTGGTTG 780
||||| ::|||
Db 204 LysGluProGlnProLysThrArgArgProTyrAspAspValValLysTrpIle 221
::||| ::|||

RESULT 11

C82436 oxygen-insensitive NAD(P)H nitroreductase VCA0637 [imported] - Vibrio cholerae (strain C/Species: Vibrio cholerae

C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C/Accession: C82436

R./Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406:477-483, 2000

A/Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A/Reference number: A82035; MUID:20406833; PMID:10952301

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-217 <HEI>

A/Cross-references: GB:AEO04393; GB:AEO03853; NID:g9658042; PIDN:AAF96539.1; GSPDB:GN000

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

C/Genetics:

A/Gene: VCAO637

A/Map position: 2

C/Superfamily: nitroreductase

Alignment Scores:	
Pred. No.: 9.13e-11	Length: 217
Score: 199.50	Matches: 55
Percent Similarity: 45.36%	Conservative: 33
Best Local Similarity: 28.35%	Mismatches: 53
Query Match: 12.80%	Indels: 53
DB: 2	Gaps: 6

US-09-647-661-1 (1-900) x C82436 (1-217)

Qy 181 CAACTATTAAACGAGCGCCATCTTGCAAGATGTTTGACGCCATATGAGTTTCTAGT 240
||| |::|
Db 5 GlnAlaSerGlnSerArgTyrSerThrLysAlaPheAspAlaSerArgLysLeuSerGlu 24
||| |::|
Qy 241 GAAGATTAGAAATAATCGCTGAATCGCCAGGCTATCGCCAAGCTCTTAGAACACGAG 300
::: ::::: ::::: |||
Db 25 GlnGlnValAlaAspLeuLysGluLeuValArgMetSerAlaSerValAsnSerGln 44
::: ::::: ::::: |||
Qy 301 CCATGGCATTTTGTGATCGTTACTAATAAGGATTTAAAAAACAATTCGATGTCACAGC 360
||| |::|
Db 45 ProTrpHisPheIleLeuAlaGlySerAspGluGlyLysThrArgIleAlaLysAlaThr 64
||| |::|
Qy 361 TAC-----TTTAATGACGAATAATGATTAAGCGCTTCACGCTTAATGCTGTA 408
||| |::|
Db 65 GlnGlyGlyPheSerPheAsnGluArgLysIleLeuAspAlaSerHisValMetValPhe 84
||| |::|
Qy 409 TGC-----TCCTTAAGACCTAGCGAGTTGTTACCA----- 438
||| |::|
Db 85 CysAlaLysThrSerIleAspAspAlaTyrLeuLeuSerLeuLeuAspAsnGluAspLys 104
||| |::|
Qy 439 -----CACGCC-----CATTCATG 453
||| |::|

Db 136 AspGlnTrpMetAlaLysGlnValTyrLeuAsnValGlyAsnPheLeuLeuGlyValGly 155
 Qy 607 TTAATGGATTGGATTCATTCATTTATGGAGGCTTTCATCCCTTAAAGTGGGTGAAGTT 666
 Db 156 AlaMetGlyLeuAspAlaValProIleGluGlyPheAsp-----AlaAlaIle 171
 Qy 667 TTAGAGAG 675
 Db 172 LeuAspGlu 174

RESULT 14

S08397
 nitroreductase (EC 1.6.6.-) - Salmonella typhimurium
 C:Species: Salmonella typhimurium
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: S08397
 R:Watanabe, M.; Ishidate Jr., M.; Nohmi, T.
 Nucleic Acids Res. 18, 1059, 1990
 A>Title: Nucleotide sequence of Salmonella typhimurium nitroreductase gene.
 A/Reference number: S08397; MUID:90192100; PMID:2179862
 A/Accession: S08397
 A/Molecule type: DNA
 A/Residues: 1-217 <WAT>
 A/Cross-references: EMBL:X17250; NID:g47792; PIDN:CAA35113.1; PID:g47793
 C:Superfamily: nitroreductase
 C:Keywords: Oxidoreductase

Alignment Scores:
 Pred. No.: 1.76e-10 Length: 217
 Score: 196.50 Matches: 52
 Percent Similarity: 48.35% Conservative: 36
 Best Local Similarity: 28.57% Mismatches: 55
 Query Match: 12.60% Indels: 39
 DB: 1 Gaps: 5

US-09-647-661-1 (1-900) x S08397 (1-217)

Qy 193 GAGCGCATCTTCAGAGATGTTTGACAGCCATTATGAGTTTCTAGTGAAGATTAGAA 252
 Db 9 GlnArgTyrSerThrLysAlaPheAspProSerLysLysLeuThrAlaGluAlaAsp 28
 Qy 253 GAATCGCTGAAATCGCCAGGCTATCGCAAGCTTCTTACACAGCCATGGCATTTT 312
 Db 29 LysIleLysThrLeuLeuGlnTyrSerProSerSerThrAsnSerGlnProTrpHisPhe.48
 Qy 313 GTGATGGTTACTAATAGGATTTAAACAAACAAATTCAGTGCACAGC----- 360
 Db 49 IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68
 Qy 361 TACTTTAATGAAGAAATGATTAAAGCGCTTCACGCTTAATGGTGGTATGCTCTTTAAGA 420
 Db 69 ThrPheAsnGluArgLysMetLeuAspAlaSerHisValValPheCysAla----- 86
 Qy 421 CCTAGCGAGTTGTTACCACAGCGCATTTACATGCAAAACCTTTTACCCGGAGTCTTAAAG 480
 Db 87 -----LysThrAlaMetAspAlaTrpLeu 95
 Qy 481 GTTAGAGTGATC-----CCTTCTTTT 501
 Db 96 GluArgValAlaAspGlnGluAspAlaAspGlyArgPheAlaThrProGluAlaLysAla 115
 Qy 502 GCTCAATGCTTGGCGTGGAGATTCACACAGCATGCAAGATTA----- 546
 Db 116 AlaAsnAspLysGlyArgPhePheAlaAspMetHisArgValSerLeuLysAspAsp 135
 Qy 547 GAAAGCTATATTTAGACCAATGCTATATCGCTGTGGGGCAAAATTTGCATGGCGGTGAGC 606
 Db 136 HisGlnTrpMetAlaLysGlnValTyrLeuAsnValGlyAsnPheLeuLeuGlyValAla 155
 Qy 607 TTAATGGATTGGATTCATTCATTTATGGAGGCTTTCATCCCTTAAAGTGGGTGAAGTT 666
 Db 156 AlaMetGlyLeuAspAlaValProIleGluGlyPheAsp-----AlaGluVal 171

Qy 667 TTAGAA 672
 Db 172 LeuAsp 173
 RESULT 15
 AD0573
 oxygen-insensitive NAD(P)H nitroreductase [imported] - Salmonella enterica subsp. enter
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A>Note: this species has also been called Salmonella typhi
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C/Accession: AD0573
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
 th, T.; Conington, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
 , S.; Moule, S.; O'Garra, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica ser
 A/Reference number: AB0502; MUID:21534947; PMID:11677608
 A/Accession: AD0573
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-217 <PAR>
 A/Cross-references: GB:AL513382; PIDN:CAD05054.1; PID:g16501830; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY0620
 C:Superfamily: nitroreductase

Alignment Scores:
 Pred. No.: 2.72e-10 Length: 217
 Score: 194.50 Matches: 51
 Percent Similarity: 48.35% Conservative: 37
 Best Local Similarity: 28.02% Mismatches: 55
 Query Match: 12.48% Indels: 39
 DB: 2 Gaps: 5

US-09-647-661-1 (1-900) x AD0573 (1-217)

Qy 193 GAGCGCATCTTCAGAGATGTTTGACAGCCATTATGAGTTTCTAGTGAAGATTAGAA 252
 Db 9 LysArgTyrSerThrLysAlaPheAspProSerLysLysLeuThrAlaGluAlaAsp 28
 Qy 253 GAATCGCTGAAATCGCCAGGCTATCGCAAGCTTCTTACACAGCCATGGCATTTT 312
 Db 29 LysValLysThrLeuLeuGlnTyrSerProSerSerThrAsnSerGlnProTrpHisPhe 48
 Qy 313 GTGATGGTTACTAATAGGATTTAAACAAACAAATTCAGTGCACAGC----- 360
 Db 49 IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68
 Qy 361 TACTTTAATGAAGAAATGATTAAAGCGCTTCAGCGTTTAATGGTGGTATGCTCTTTAAGA 420
 Db 69 ThrPheAsnGluArgLysMetLeuAspAlaSerHisValValPheCysAla----- 86
 Qy 421 CCTAGCGAGTTGTTACCACAGCGCATTTACATGCAAAACCTTTTACCCGGAGTCTTAAAG 480
 Db 87 -----LysThrAlaMetAspAlaTrpLeu 95
 Qy 481 GTTAGAGTGATC-----CCTTCTTTT 501
 Db 96 GluArgValAlaAspGlnGluAspAlaAspGlyArgPheAlaThrProGluAlaLysAla 115
 Qy 502 GCTCAATGCTTGGCGTGGAGATTCACACAGCATGCAAGATTA----- 546
 Db 116 AlaAsnAspLysGlyArgPhePheAlaAspMetHisArgValSerLeuLysAspAsp 135
 Qy 547 GAAAGCTATATTTAGACCAATGCTATATCGCTGTGGGGCAAAATTTGCATGGCGGTGAGC 606
 Db 136 HisGlnTrpMetAlaLysGlnValTyrLeuAsnValGlyAsnPheLeuLeuGlyValAla 155
 Qy 607 TTAATGGATTGGATTCATTCATTTATGGAGGCTTTCATCCCTTAAAGTGGGTGAAGTT 666
 Db 156 AlaMetGlyLeuAspAlaValProIleGluGlyPheAsp-----AlaGluVal 171

Qy 667 TTAGAA 672
| | | | |
Db 172 LeuAsp 173

Search completed: April 20, 2004, 19:29:01
Job time : 28.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 20, 2004, 19:12:57 ; Search time 15 Seconds

(without alignments)
6248.419 Million cell updates/sec

Title: US-09-647-661-1
Perfect score: 1559
Sequence: 1 tcagagattttacagagc.....gggtgttttcaaggtttc 900

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_epool_p/US09647661/runat_20042004_135347_2651/app_query.fasta_1.1095
-DB=SwissProt_42-QFMT=fastan -SUPFIX=rsp -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=numat40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HRAPIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09647661@cgn 1 1 16 @runat_20042004_135347_2651 -NCFU=3
-NO MMAP -LARGEOBURY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	246	15.8	284	1 LGT_HELPY	Q25609 helicobacte
2	229	14.7	283	1 LGT_HELPY	Q82k06 helicobacte
3	223.5	14.3	221	1 YFAC_BACSU	Q34475 bacillus su
4	218.5	14.0	220	1 YC78_HAEIN	Q57431 haemophilus
5	198.5	12.7	217	1 NFNB_ENTCL	Q01234 enterobacte
6	196.5	12.6	217	1 NFNB_SALTY	P15888 salmorella
7	188.5	12.1	217	1 NFNB_ECOLI	P38489 escherichia
8	185.5	11.9	201	1 V12C_BACSU	P81102 bacillus su
9	172.5	11.1	209	1 YGCI_BACSU	P96707 bacillus su
10	169.5	10.9	210	1 DRGA_SYNT3	Q55233 synecocyst
11	154	9.9	217	1 FRAL_VIBET	P46072 vibrio fisc
12	153	9.8	174	1 YM67_ARCFU	Q28017 archaeoglob
13	137.5	8.8	205	1 YDFN_BACSU	P96692 bacillus su
14	136.5	8.8	205	1 NOX_TRETH	Q60049 thermus the
15	126	8.1	191	1 Y120_METTH	O26223 methanobact
16	124	8.0	172	1 Y236_ARCFU	Q30013 archaeoglob
17	103.5	6.6	359	1 YK45_CABEL	P34273 caenorhabdi
18	97.5	6.3	274	1 DAPD_BUCAL	P57323 buchnera ap

19	91	5.8	1125	1 MFD_BORBU	O51568 borrelia bu
20	90.5	5.8	281	1 LGT_BUCAL	P57514 buchnera ap
21	90.5	5.8	353	1 Y772_METIA	Q58182 methanococ
22	89	5.7	1001	1 MFD_HELPY	Q2zj57 helicobacte
23	88	5.6	184	1 YDJA_HAEIN	P45244 haemophilus
24	87.5	5.6	202	1 YZEL_STRCO	Q82bvo streptomyce
25	87.5	5.6	273	1 DAPD_BUCAP	O85290 buchnera ap
26	87.5	5.6	291	1 LGT_ECOLI	P37149 escherichia
27	87.5	5.6	901	1 MALT_ECOLI	P06993 escherichia
28	86.5	5.5	196	1 Y096_AGRTS	P58792 agrobacteri
29	86.5	5.5	999	1 MFD_HELPY	O86066 helicobacte
30	86	5.7	315	1 OSAY_HUMAN	Q8ngz2 homo sapien
31	85.5	5.5	459	1 MURD_LACEL	Q88v80 lactobacill
32	85	5.5	484	1 SYE_MYCGE	P47700 mycoplasma
33	85	5.5	901	1 R8L3_ARATH	Q8fjb5 arabidopsis
34	84.5	5.4	245	1 KDSB_FUSNN	O8fia8 fusobacteri
35	84	5.4	452	1 DHAP_RAT	P11883 rattus norv
36	84	5.4	1102	1 MISC_CHICK	P29616 gallus gall
37	83.5	5.4	207	1 RFOA_EUGAN	Q8e194 euglena ana
38	83.5	5.4	1581	1 ACC8_CRICR	Q09427 cricetus cr
39	83	5.3	249	1 NFRA_BACSV	P39605 bacillus su
40	83	5.3	472	1 PAP1_FOWPV	Q9j5b8 fowlpox vir
41	83	5.3	535	1 Y295_STR3A	O86762 streptococc
42	83	5.3	888	1 LONI_ARATH	O64948 arabidopsis
43	83	5.3	1276	1 BXA_CLOBO	P13321 clostridium
44	82.5	5.3	751	1 TALA_POVHA	P03075 hamster pol
45	82.5	5.3	1161	1 BM2K_HUMAN	Q9msy1 homo sapien

ALIGNMENTS

RESULT 1
LGT_HELPY

ID	LGT_HELPY	STANDARD;	PRT;	284 AA.
AC	Q25609; Q30584; Q30586;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Fropiloprotein diacylglycerol transferase (EC 2.4.99.-).			
GN	LGT OR HP0955.			
OS	Helicobacter pylori (Campylobacter pylori).			
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;			
OC	Helicobacteraceae; Helicobacter.			
OX	NCBI_TaxID=210;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=26695 / ATCC 700392;			
RX	MEDLINE=97394467; PubMed=9252185;			
RA	Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,			
RA	Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,			
RA	Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,			
RA	Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,			
RA	McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,			
RA	Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,			
RA	Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,			
RA	Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,			
RA	Venter J.C.;			
RA	"The complete genome sequence of the gastric pathogen Helicobacter			
RT	pylori.";			
RL	Nature 368:539-547(1997).			
RP	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=HP439, and HP500;			
RA	Goodwin A., Berg D.E., Hoffman P.S.;			
RL	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.			
CC	-/- FUNCTION: Transfers the N-acyl diglyceride group on what will			
CC	become the N-terminal cysteine of membrane lipoproteins (By			
CC	similarity).			
CC	-/- PATHWAY: Lipoproteins biosynthesis; first step.			
CC	-/- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane			
CC	(Potential).			
CC	-/- SIMILARITY: Belongs to the lgt family.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AB000604; AAC07998.1; -
 CC EMBL; AF012552; AAC46348.1; -
 CC EMBL; AF012553; AAC46350.1; -
 CC PIR; C64639; C64639.
 CC TIGR; HP0955; -
 CC InterPro: IPR001640; LGT.
 CC Pfam; PF01790; LGT; 1.
 CC ProDom; PD005412; LGT; 1.
 CC TIGRfam; TIGR00544; LGT; 1.
 CC PROSITE; PS01311; LGT; 1.
 CC TRANSFERASE; Transmembrane; Inner membrane; Complete proteome.
 CC TRANSMEM 14 34
 CC TRANSMEM 62 82
 CC TRANSMEM 106 126
 CC TRANSMEM 136 156
 CC TRANSMEM 190 210
 CC TRANSMEM 218 238
 CC TRANSMEM 252 272
 CC TRANSMEM 284 326
 CC TRANSMEM 326 364
 CC TRANSMEM 364 402
 CC TRANSMEM 402 440
 CC TRANSMEM 440 478
 CC TRANSMEM 478 516
 CC TRANSMEM 516 554
 CC TRANSMEM 554 592
 CC TRANSMEM 592 630
 CC TRANSMEM 630 668
 CC TRANSMEM 668 706
 CC TRANSMEM 706 744
 CC TRANSMEM 744 782
 CC TRANSMEM 782 820
 CC TRANSMEM 820 858
 CC TRANSMEM 858 896
 CC TRANSMEM 896 934
 CC TRANSMEM 934 972
 CC TRANSMEM 972 1010
 CC TRANSMEM 1010 1048
 CC TRANSMEM 1048 1086
 CC TRANSMEM 1086 1124
 CC TRANSMEM 1124 1162
 CC TRANSMEM 1162 1200
 CC TRANSMEM 1200 1238
 CC TRANSMEM 1238 1276
 CC TRANSMEM 1276 1314
 CC TRANSMEM 1314 1352
 CC TRANSMEM 1352 1390
 CC TRANSMEM 1390 1428
 CC TRANSMEM 1428 1466
 CC TRANSMEM 1466 1504
 CC TRANSMEM 1504 1542
 CC TRANSMEM 1542 1580
 CC TRANSMEM 1580 1618
 CC TRANSMEM 1618 1656
 CC TRANSMEM 1656 1694
 CC TRANSMEM 1694 1732
 CC TRANSMEM 1732 1770
 CC TRANSMEM 1770 1808
 CC TRANSMEM 1808 1846
 CC TRANSMEM 1846 1884
 CC TRANSMEM 1884 1922
 CC TRANSMEM 1922 1960
 CC TRANSMEM 1960 1998
 CC TRANSMEM 1998 2036
 CC TRANSMEM 2036 2074
 CC TRANSMEM 2074 2112
 CC TRANSMEM 2112 2150
 CC TRANSMEM 2150 2188
 CC TRANSMEM 2188 2226
 CC TRANSMEM 2226 2264
 CC TRANSMEM 2264 2302
 CC TRANSMEM 2302 2340
 CC TRANSMEM 2340 2378
 CC TRANSMEM 2378 2416
 CC TRANSMEM 2416 2454
 CC TRANSMEM 2454 2492
 CC TRANSMEM 2492 2530
 CC TRANSMEM 2530 2568
 CC TRANSMEM 2568 2606
 CC TRANSMEM 2606 2644
 CC TRANSMEM 2644 2682
 CC TRANSMEM 2682 2720
 CC TRANSMEM 2720 2758
 CC TRANSMEM 2758 2796
 CC TRANSMEM 2796 2834
 CC TRANSMEM 2834 2872
 CC TRANSMEM 2872 2910
 CC TRANSMEM 2910 2948
 CC TRANSMEM 2948 2986
 CC TRANSMEM 2986 3024
 CC TRANSMEM 3024 3062
 CC TRANSMEM 3062 3100
 CC TRANSMEM 3100 3138
 CC TRANSMEM 3138 3176
 CC TRANSMEM 3176 3214
 CC TRANSMEM 3214 3252
 CC TRANSMEM 3252 3290
 CC TRANSMEM 3290 3328
 CC TRANSMEM 3328 3366
 CC TRANSMEM 3366 3404
 CC TRANSMEM 3404 3442
 CC TRANSMEM 3442 3480
 CC TRANSMEM 3480 3518
 CC TRANSMEM 3518 3556
 CC TRANSMEM 3556 3594
 CC TRANSMEM 3594 3632
 CC TRANSMEM 3632 3670
 CC TRANSMEM 3670 3708
 CC TRANSMEM 3708 3746
 CC TRANSMEM 3746 3784
 CC TRANSMEM 3784 3822
 CC TRANSMEM 3822 3860
 CC TRANSMEM 3860 3898
 CC TRANSMEM 3898 3936
 CC TRANSMEM 3936 3974
 CC TRANSMEM 3974 4012
 CC TRANSMEM 4012 4050
 CC TRANSMEM 4050 4088
 CC TRANSMEM 4088 4126
 CC TRANSMEM 4126 4164
 CC TRANSMEM 4164 4202
 CC TRANSMEM 4202 4240
 CC TRANSMEM 4240 4278
 CC TRANSMEM 4278 4316
 CC TRANSMEM 4316 4354
 CC TRANSMEM 4354 4392
 CC TRANSMEM 4392 4430
 CC TRANSMEM 4430 4468
 CC TRANSMEM 4468 4506
 CC TRANSMEM 4506 4544
 CC TRANSMEM 4544 4582
 CC TRANSMEM 4582 4620
 CC TRANSMEM 4620 4658
 CC TRANSMEM 4658 4696
 CC TRANSMEM 4696 4734
 CC TRANSMEM 4734 4772
 CC TRANSMEM 4772 4810
 CC TRANSMEM 4810 4848
 CC TRANSMEM 4848 4886
 CC TRANSMEM 4886 4924
 CC TRANSMEM 4924 4962
 CC TRANSMEM 4962 5000
 CC TRANSMEM 5000 5038
 CC TRANSMEM 5038 5076
 CC TRANSMEM 5076 5114
 CC TRANSMEM 5114 5152
 CC TRANSMEM 5152 5190
 CC TRANSMEM 5190 5228
 CC TRANSMEM 5228 5266
 CC TRANSMEM 5266 5304
 CC TRANSMEM 5304 5342
 CC TRANSMEM 5342 5380
 CC TRANSMEM 5380 5418
 CC TRANSMEM 5418 5456
 CC TRANSMEM 5456 5494
 CC TRANSMEM 5494 5532
 CC TRANSMEM 5532 5570
 CC TRANSMEM 5570 5608
 CC TRANSMEM 5608 5646
 CC TRANSMEM 5646 5684
 CC TRANSMEM 5684 5722
 CC TRANSMEM 5722 5760
 CC TRANSMEM 5760 5798
 CC TRANSMEM 5798 5836
 CC TRANSMEM 5836 5874
 CC TRANSMEM 5874 5912
 CC TRANSMEM 5912 5950
 CC TRANSMEM 5950 5988
 CC TRANSMEM 5988 6026
 CC TRANSMEM 6026 6064
 CC TRANSMEM 6064 6102
 CC TRANSMEM 6102 6140
 CC TRANSMEM 6140 6178
 CC TRANSMEM 6178 6216
 CC TRANSMEM 6216 6254
 CC TRANSMEM 6254 6292
 CC TRANSMEM 6292 6330
 CC TRANSMEM 6330 6368
 CC TRANSMEM 6368 6406
 CC TRANSMEM 6406 6444
 CC TRANSMEM 6444 6482
 CC TRANSMEM 6482 6520
 CC TRANSMEM 6520 6558
 CC TRANSMEM 6558 6596
 CC TRANSMEM 6596 6634
 CC TRANSMEM 6634 6672
 CC TRANSMEM 6672 6710
 CC TRANSMEM 6710 6748
 CC TRANSMEM 6748 6786
 CC TRANSMEM 6786 6824
 CC TRANSMEM 6824 6862
 CC TRANSMEM 6862 6900
 CC TRANSMEM 6900 6938
 CC TRANSMEM 6938 6976
 CC TRANSMEM 6976 7014
 CC TRANSMEM 7014 7052
 CC TRANSMEM 7052 7090
 CC TRANSMEM 7090 7128
 CC TRANSMEM 7128 7166
 CC TRANSMEM 7166 7204
 CC TRANSMEM 7204 7242
 CC TRANSMEM 7242 7280
 CC TRANSMEM 7280 7318
 CC TRANSMEM 7318 7356
 CC TRANSMEM 7356 7394
 CC TRANSMEM 7394 7432
 CC TRANSMEM 7432 7470
 CC TRANSMEM 7470 7508
 CC TRANSMEM 7508 7546
 CC TRANSMEM 7546 7584
 CC TRANSMEM 7584 7622
 CC TRANSMEM 7622 7660
 CC TRANSMEM 7660 7698
 CC TRANSMEM 7698 7736
 CC TRANSMEM 7736 7774
 CC TRANSMEM 7774 7812
 CC TRANSMEM 7812 7850
 CC TRANSMEM 7850 7888
 CC TRANSMEM 7888 7926
 CC TRANSMEM 7926 7964
 CC TRANSMEM 7964 8002
 CC TRANSMEM 8002 8040
 CC TRANSMEM 8040 8078
 CC TRANSMEM 8078 8116
 CC TRANSMEM 8116 8154
 CC TRANSMEM 8154 8192
 CC TRANSMEM 8192 8230
 CC TRANSMEM 8230 8268
 CC TRANSMEM 8268 8306
 CC TRANSMEM 8306 8344
 CC TRANSMEM 8344 8382
 CC TRANSMEM 8382 8420
 CC TRANSMEM 8420 8458
 CC TRANSMEM 8458 8496
 CC TRANSMEM 8496 8534
 CC TRANSMEM 8534 8572
 CC TRANSMEM 8572 8610
 CC TRANSMEM 8610 8648
 CC TRANSMEM 8648 8686
 CC TRANSMEM 8686 8724
 CC TRANSMEM 8724 8762
 CC TRANSMEM 8762 8800
 CC TRANSMEM 8800 8838
 CC TRANSMEM 8838 8876
 CC TRANSMEM 8876 8914
 CC TRANSMEM 8914 8952
 CC TRANSMEM 8952 8990
 CC TRANSMEM 8990 9028
 CC TRANSMEM 9028 9066
 CC TRANSMEM 9066 9104
 CC TRANSMEM 9104 9142
 CC TRANSMEM 9142 9180
 CC TRANSMEM 9180 9218
 CC TRANSMEM 9218 9256
 CC TRANSMEM 9256 9294
 CC TRANSMEM 9294 9332
 CC TRANSMEM 9332 9370
 CC TRANSMEM 9370 9408
 CC TRANSMEM 9408 9446
 CC TRANSMEM 9446 9484
 CC TRANSMEM 9484 9522
 CC TRANSMEM 9522 9560
 CC TRANSMEM 9560 9598
 CC TRANSMEM 9598 9636
 CC TRANSMEM 9636 9674
 CC TRANSMEM 9674 9712
 CC TRANSMEM 9712 9750
 CC TRANSMEM 9750 9788
 CC TRANSMEM 9788 9826
 CC TRANSMEM 9826 9864
 CC TRANSMEM 9864 9902
 CC TRANSMEM 9902 9940
 CC TRANSMEM 9940 9978
 CC TRANSMEM 9978 10016
 CC TRANSMEM 10016 10054
 CC TRANSMEM 10054 10092
 CC TRANSMEM 10092 10130
 CC TRANSMEM 10130 10168
 CC TRANSMEM 10168 10206
 CC TRANSMEM 10206 10244
 CC TRANSMEM 10244 10282
 CC TRANSMEM 10282 10320
 CC TRANSMEM 10320 10358
 CC TRANSMEM 10358 10396
 CC TRANSMEM 10396 10434
 CC TRANSMEM 10434 10472
 CC TRANSMEM 10472 10510
 CC TRANSMEM 10510 10548
 CC TRANSMEM 10548 10586
 CC TRANSMEM 10586 10624
 CC TRANSMEM 10624 10662
 CC TRANSMEM 10662 10700
 CC TRANSMEM 10700 10738
 CC TRANSMEM 10738 10776
 CC TRANSMEM 10776 10814
 CC TRANSMEM 10814 10852
 CC TRANSMEM 10852 10890
 CC TRANSMEM 10890 10928
 CC TRANSMEM 10928 10966
 CC TRANSMEM 10966 11004
 CC TRANSMEM 11004 11042
 CC TRANSMEM 11042 11080
 CC TRANSMEM 11080 11118
 CC TRANSMEM 11118 11156
 CC TRANSMEM 11156 11194
 CC TRANSMEM 11194 11232
 CC TRANSMEM 11232 11270
 CC TRANSMEM 11270 11308
 CC TRANSMEM 11308 11346
 CC TRANSMEM 11346 11384
 CC TRANSMEM 11384 11422
 CC TRANSMEM 11422 11460
 CC TRANSMEM 11460 11498
 CC TRANSMEM 11498 11536
 CC TRANSMEM 11536 11574
 CC TRANSMEM 11574 11612
 CC TRANSMEM 11612 11650
 CC TRANSMEM 11650 11688
 CC TRANSMEM 11688 11726
 CC TRANSMEM 11726 11764
 CC TRANSMEM 11764 11802
 CC TRANSMEM 11802 11840
 CC TRANSMEM 11840 11878
 CC TRANSMEM 11878 11916
 CC TRANSMEM 11916 11954
 CC TRANSMEM 11954 11992
 CC TRANSMEM 11992 12030
 CC TRANSMEM 12030 12068
 CC TRANSMEM 12068 12106
 CC TRANSMEM 12106 12144
 CC TRANSMEM 12144 12182
 CC TRANSMEM 12182 12220
 CC TRANSMEM 12220 12258
 CC TRANSMEM 12258 12296
 CC TRANSMEM 12296 12334
 CC TRANSMEM 12334 12372
 CC TRANSMEM 12372 12410
 CC TRANSMEM 12410 12448
 CC TRANSMEM 12448 12486
 CC TRANSMEM 12486 12524
 CC TRANSMEM 12524 12562
 CC TRANSMEM 12562 12600
 CC TRANSMEM 12600 12638
 CC TRANSMEM 12638 12676
 CC TRANSMEM 12676 12714
 CC TRANSMEM 12714 12752
 CC TRANSMEM 12752 12790
 CC TRANSMEM 12790 12828
 CC TRANSMEM 12828 12866
 CC TRANSMEM 12866 12904
 CC TRANSMEM 12904 12942
 CC TRANSMEM 12942 12980
 CC TRANSMEM 12980 13018
 CC TRANSMEM 13018 13056
 CC TRANSMEM 13056 13094
 CC TRANSMEM 13094 13132
 CC TRANSMEM 13132 13170
 CC TRANSMEM 13170 13208
 CC TRANSMEM 13208 13246
 CC TRANSMEM 13246 13284
 CC TRANSMEM 13284 13322
 CC TRANSMEM 13322 13360
 CC TRANSMEM 13360 13398
 CC TRANSMEM 13398 13436
 CC TRANSMEM 13436 13474
 CC TRANSMEM 13474 13512
 CC TRANSMEM 13512 13550
 CC TRANSMEM 13550 13588
 CC TRANSMEM 13588 13626
 CC TRANSMEM 13626 13664
 CC TRANSMEM 13664 13702
 CC TRANSMEM 13702 13740
 CC TRANSMEM 13740 13778
 CC TRANSMEM 13778 13816
 CC TRANSMEM 13816 13854
 CC TRANSMEM 13854 13892
 CC TRANSMEM 13892 13930
 CC TRANSMEM 13930 13968
 CC TRANSMEM 13968 14006
 CC TRANSMEM 14006 14044
 CC TRANSMEM 14044 14082
 CC TRANSMEM 14082 14120
 CC TRANSMEM 14120 14158
 CC TRANSMEM 14158 14196
 CC TRANSMEM 14196 14234
 CC TRANSMEM 14234 14272
 CC TRANSMEM 14272 14310
 CC TRANSMEM 14310 14348
 CC TRANSMEM 14348 14386
 CC TRANSMEM 14386 14424
 CC TRANSMEM 14424 14462
 CC TRANSMEM 14462 14500
 CC TRANSMEM 14500 14538
 CC TRANSMEM 14538 14576
 CC TRANSMEM 14576 14614
 CC TRANSMEM 14614 14652
 CC TRANSMEM 14652 14690
 CC TRANSMEM 14690 14728
 CC TRANSMEM 14728 14766
 CC TRANSMEM 14766 14804
 CC TRANSMEM 14804 14842
 CC TRANSMEM 14842 14880
 CC TRANSMEM 14880 14918
 CC TRANSMEM 14918 14956
 CC TRANSMEM 14956 14994
 CC TRANSMEM 14994 15032
 CC TRANSMEM 15032 15070
 CC TRANSMEM 15070 15108
 CC TRANSMEM 15108 15146
 CC TRANSMEM 15146 15184
 CC TRANSMEM 15184 15222
 CC TRANSMEM 15222 15260
 CC TRANSMEM 15260 15298
 CC TRANSMEM 15298 15336
 CC TRANSMEM 15336 15374
 CC TRANSMEM 15374 15412
 CC TRANSMEM 15412 15450
 CC TRANSMEM 15450 15488
 CC TRANSMEM 15488 15526
 CC TRANSMEM 15526 15564
 CC TRANSMEM 15564 15602
 CC TRANSMEM 15602 15640
 CC TRANSMEM 15640 15678
 CC TRANSMEM 15678 15716
 CC TRANSMEM 15716 15754
 CC TRANSMEM 15754 15792
 CC TRANSMEM 15792 15830
 CC TRANSMEM 15830 15868
 CC TRANSMEM 15868 15906
 CC TRANSMEM 15906 15944
 CC TRANSMEM 15944 15982
 CC TRANSMEM 15982 16020
 CC TRANSMEM 16020 16058
 CC TRANSMEM 16058 16096
 CC TRANSMEM 16096 16134
 CC TRANSMEM 16134 16172
 CC TRANSMEM 16172 16210
 CC TRANSMEM 16210 16248
 CC TRANSMEM 16248 16286
 CC TRANSMEM 16286 16324
 CC TRANSMEM 16324 16362
 CC TRANSMEM 16362 16400
 CC TRANSMEM 16400 16438
 CC TRANSMEM 16438 16476
 CC TRANSMEM 16476 16514
 CC TRANSMEM 16514 16552
 CC TRANSMEM 16552 16590
 CC TRANSMEM 16590 16628
 CC TRANSMEM 16628 16666
 CC TRANSMEM 16666 16704
 CC TRANSMEM 16704 16742
 CC TRANSMEM 16742 16780
 CC TRANSMEM 16780 16818
 CC TRANSMEM 16818 16856
 CC TRANSMEM 16856 16894
 CC TRANSMEM 16894 16932
 CC TRANSMEM 16932 16970
 CC TRANSMEM 16970 17008
 CC TRANSMEM 17008 17046
 CC TRANSMEM 17046 17084
 CC TRANSMEM 17084 17122
 CC TRANSMEM 17122 17160
 CC TRANSMEM 17160 17198
 CC TRANSMEM 17198 17236
 CC TRANSMEM 17236 17274
 CC TRANSMEM 17274 17312
 CC TRANSMEM 17312 17350
 CC TRANSMEM 17350 17388
 CC TRANSMEM 17388 17426
 CC TRANSMEM 17426 17464
 CC TRANSMEM 17464 17502
 CC TRANSMEM 17502 17540
 CC TRANSMEM 17540 17578
 CC TRANSMEM 17578 17616
 CC TRANSMEM 17616 17654
 CC TRANSMEM 17654 17692
 CC TRANSMEM 17692 17730
 CC TRANSMEM 17730 17768
 CC TRANSMEM 17768 17806
 CC TRANSMEM 17806 17844
 CC TRANSMEM 17844 17882
 CC TRANSMEM 17882 17920
 CC TRANSMEM 17920 17958
 CC TRANSMEM 17958 17996
 CC TRANSMEM 17996 18034
 CC TRANSMEM 18034 18072
 CC TRANSMEM 18072 18110
 CC TRANSMEM 18110 18148
 CC TRANSMEM 18148 18186
 CC TRANSMEM 18186 18224
 CC TRANSMEM 18224 18262
 CC TRANSMEM 18262 18300
 CC TRANSMEM 18300 18338
 CC TRANSMEM 18338 18376
 CC TRANSMEM 18376 18414
 CC TRANSMEM 18414 18452
 CC TRANSMEM 18452 18490
 CC TRANSMEM 18490 18528
 CC TRANSMEM 18528 18566
 CC TRANSMEM 18566 18604
 CC TRANSMEM 18604 18642
 CC TRANSMEM 18642 18680
 CC TRANSMEM 18680 18718
 CC TRANSMEM 18718 18756
 CC TRANSMEM 18756 18794
 CC TRANSMEM 18794 18832
 CC TRANSMEM 18832 18870
 CC TRANSMEM 18870 18908
 CC TRANSMEM 18908 18946
 CC TRANSMEM 18946 18984
 CC TRANSMEM 18984 19022
 CC TRANSMEM 19022 19060
 CC TRANSMEM 19060 19098
 CC TRANSMEM 19098 19136
 CC TRANSMEM 19136 19174
 CC TRANSMEM 19174 19212
 CC TRANSMEM 19212 19250
 CC TRANSMEM 19250 19288
 CC TRANSMEM 19288 19326
 CC TRANSMEM 19326 19364
 CC TRANSMEM 19364 19402
 CC TRANSMEM 19402 19440
 CC TRANSMEM 19440 19478
 CC TRANSMEM 19478 19516
 CC TRANSMEM 19516 19554
 CC TRANSMEM 19554 19592
 CC TRANSMEM 19592 19630
 CC TRANSMEM 19630 19668
 CC TRANSMEM 19668 19706
 CC TRANSMEM 19706 19744
 CC TRANSMEM 19744 19782
 CC TRANSMEM 19782 19820
 CC TRANSMEM 19820 19858
 CC TRANSMEM 19858 19896
 CC TRANSMEM 19896 19934
 CC TRANSMEM 19934 19972
 CC TRANSMEM 19972 20010
 CC TRANSMEM 20010 20048
 CC TRANSMEM 20048 20086
 CC TRANSMEM 20086 20124
 CC TRANSMEM 20124 20162
 CC TRANSMEM 20162 20200
 CC TRANSMEM 20200 20238
 CC TRANSMEM 20238 20276
 CC TRANSMEM 20276 20314
 CC TRANSMEM 20314 20352
 CC TRANSMEM 20352 20390
 CC TRANSMEM 20390 20428
 CC TRANSMEM 20428 20466
 CC TRANSMEM 20466 20504
 CC TRANSMEM 20504 20542
 CC TRANSMEM 20542 20580
 CC TRANSMEM 20580 20618
 CC TRANSMEM 20618 20656
 CC TRANSMEM 20656 20694
 CC TRANSMEM 20694 20732
 CC TRANSMEM 20732 20770
 CC TRANSMEM 20770 20808
 CC TRANSMEM 20808 20846
 CC TRANSMEM 20846 20884
 CC TRANSMEM 20884 20922
 CC TRANSMEM 20922 20960
 CC TRANSMEM 20960 21000

US-09-647-661-1 (1-900) x LGT_HELPJ (1-283)

QY 2 GCAGATTTTACAGAGCCAGATAGCCAAATGGGGTTATTTTAAATTTGAGCATG 61
 Db 235 AlaGluPheTyArgGluProAspSerGlnLeuGlyValTyPheLeuAsnLeuSerMet 254
 QY 62 GGGCAGATTTTAAAGCTTATTTATGTTAGTTAGTTTTCATTAGGATTTTATTTATCTACA 121
 Db 255 GlyGlnIleLeuSerValPheMetValIleValSerLeuGlyIleLeuLeuTyAlaThr 274
 QY 122 AAAAATCTTAAATAAAGGAA 145
 Db 275 LysAsnSerLysLysLysLysGlu 282
 RESULT 3
 YFKO_BACSU STANDARD; PRT; 221 AA.
 ID_YFKO_BACSU
 AC 034475;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Putative NAD(P)H nitroreductase yfko (EC 1.1.1.1).
 GN YFKO OR BSU07830.

Alignment Scores:
 Pred. No.: 6,07e-14 Length: 283
 Score: 229.00 Matches: 45
 Percent Similarity: 100.00% Conservative: 3
 Best Local Similarity: 93.75% Mismatches: 0
 Query Match: 14.69% Indels: 0
 DB: 1 Gaps: 0

US-09-647-661-1 (1-900) x LGT_HELPJ (1-284)

QY 2 GCAGATTTTACAGAGCCAGATAGCCAAATGGGGTTATTTTAAATTTGAGCATG 61
 Db 235 AlaGluPheTyArgGluProAspSerGlnMetGlyValTyPheLeuAsnLeuSerMet 254
 QY 62 GGGCAGATTTTAAAGCTTATTTATGTTAGTTAGTTTTCATTAGGATTTTATTTATCTACA 121
 Db 255 GlyGlnIleLeuSerPheMetValIleValSerLeuGlyIleLeuLeuTyAlaThr 274
 QY 122 AAAAATCTTAAATAAAGGAAATCAA 151
 Db 275 LysAsnSerLysLysLysLysGluAsnGln 284
 RESULT 2
 LGT_HELPJ
 ID_YFKO_HELPJ STANDARD; PRT; 283 AA.
 AC 092KPF;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Prolipoprotein diacylglycerol transferase (EC 2.4.99.-).
 GN LGT OR JHP089.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]
 RP SEQUENCE FROM N.A.

[illegible]

Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.",
 RL Science 269:496-512(1995).
 RN [2]
 RP IDENTIFICATION BY MASS SPECTROMETRY.
 RX MEDLINE=20137488; PubMed=10675023;
 RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
 Gray C., Fountoulakis M.,
 RA "Two-dimensional map of the proteome of Haemophilus influenzae";
 RL Electrophoresis 21:411-429(2000).
 CC -1- COFACTOR: FMN (Potential).
 CC -1- SIMILARITY: Belongs to the nitroreductase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U32807; AAC22926.1; -
 DR PIR; B64114; B64114.
 DR HSP; P46072; 1VFR.
 DR TIGR; H11278; -
 DR InterPro: IPR000415; Nitroreductase.
 DR Pfam; PF00881; Nitroreductase, 1.
 DR Oxidoreductase; NAD; NADP; Flavoprotein; FMN; Complete proteome.
 KW NP_BIND 155 160 NAD OR NADP (BY SIMILARITY).
 FT SEQUENCE 220 AA; 25189 MW; AB0D804601BC975E CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 6.03e-13 Length: 220
 Score: 218.50 Matches: 59
 Percent Similarity: 48.89% Conservative: 51
 Best Local Similarity: 26.22% Mismatches: 94
 Query Match: 14.02% Indels: 21
 DB: 1 Gaps: 4

US-09-647-661-1 (1-900) x YC78_HABIN (1-220)

QY 151 ATGAATTTTGGATCAGGAAAAGAACACTATTAAACGAGCGCCATTCTTGCAG 210
 DB 1 MetThrGlnLeuThrArgGluGlnValLeuGluLeuPheHisGlnArgSerThrArg 20
 QY 211 ATGTTTACACCCATTATGATTTCTAGTGAAGAAATAGAGAAATCGCTGAATCGCC 270
 DB 21 TyrTyrAspProThrLysLysLysLysLysLysLysLysLysLysLysLysLys 40
 QY 271 AGGTATCGCAAGCTCTTCAACACGACGACGATGCGATTGTGTGATGCTTAAATAG 330
 DB 41 ArgLeuSerProSerSerValGlySerGluProThrLysPheLeuValLysLysLys 60
 QY 331 GATTAAAAACAAATTCAGTCACGCTACCTTTAATGAAGAAATGATTAAGCGCT 390
 DB 61 ThrLeuArgGluLysMetLysProPheSerIrpGlyMetLeuGlnLeuAspAsnLys 80
 QY 391 TCAGGTTTAATGGTGTATGCTCTTAAAGACCTAGCGATTGTGTACACAGCGCCATTAC 450
 DB 81 SerHisLeuValValLysLeuAlaLysLysLysLysLysLysLysLysLysLysLys 94
 QY 451 ATGCAAAACCTTTACCGGAGCTCTTATAGGTAGATGATCCCTCTTTGCTCAATG 510
 DB 95 AspSerProPheValAspValMetAlaArgLysGlyLeuAsnAlaGluGlnGln 114
 QY 511 CTGGCGTGAATTCACACGACGATGCAAGAGATAGAAAGCTATATTTTAGAG----- 564
 DB 115 AlaAlaLeuThrLysTyrLysAlaLeuGlnGluAspMetLysLeuLeuGluAsnAsp 134
 QY 565 -----CAATGCTATATCGCTGTGGCGCAAAATTGCGATG 597
 DB 135 ArgThrLeuPheAspTyrCysSerLysGlnThrTyrIleAlaLeuAlaAsnMetLeuThr 154

QY 598 GCGTGAGCTTAATGGATTGGATAGTATGTCATTATGAGCGTTTGATCCTTTAAAGTG 657
 DB 155 GlyAlaSerAlaLeuGlyIleAspSerCysProIleGluGlyPheHisTyrAspLysMet 174
 QY 658 GGTGAAGTTTAA---GAGAGCGTATCAATAAGCCTAAA-----ATGCGCATGCTTG 705
 DB 175 AsnGluCysLeuAlaGluGluGlyLeuPheAspProGlnGlnTyrAlaValSerValAla 194
 QY 706 ATCGCTTTGGCAAGAGGGTGGCAGACGAGCCAGCAAAATCAAGAAATCAAAAGTTGAT 765
 DB 195 AlaThrPheGlyTyrArgSerArgAspIleAlaLysLysSerArgLysGlyLeuAspGlu 214
 QY 766 GCGATTACTTGGTTG 780
 DB 215 ValValLysIrpVal 219
 RESULT 5
 ID_NFN ENTCL STANDARD; PRT; 217 AA.
 AC Q01234;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Oxygen-insensitive NAD(P)H nitroreductase (EC 1.-.-.-).
 GN NFN OR NFS1.
 OS Enterobacter cloacae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Enterobacter.
 OX NCBI_TaxID=550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 43560 / 96-3;
 RX MEDLINE=91154203; PubMed=1999406;
 RA Bryant C., Hubbard L., McElroy W.D.;
 RT "Cloning, nucleotide sequence, and expression of the nitroreductase
 gene from Enterobacter cloacae.";
 RL J. Biol. Chem. 266:4126-4130(1991).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=91154202; PubMed=1999405;
 RA Bryant C., Deluca M.;
 RT "Purification and characterization of an oxygen-insensitive NAD(P)H
 nitroreductase from Enterobacter cloacae.";
 RL J. Biol. Chem. 266:4119-4125(1991).
 CC -1- FUNCTION: REDUCTION OF A VARIETY OF NITROAROMATIC COMPOUNDS USING
 NADH (AND TO LESSER EXTENT NADPH) AS SOURCE OF REDUCING
 EQUIVALENTS; TWO ELECTRONS ARE TRANSFERRED.
 CC -1- COFACTOR: FMN.
 CC -1- PATHWAY: THE NITROREDUCTASE MIGHT BE INVOLVED IN THE QUINONE
 METABOLISM. IT SHOWS FUNCTIONAL SIMILARITY TO MAMMALIAN
 QUINONE REDUCTASES.
 CC -1- SUBUNIT: Monomer.
 CC -1- SIMILARITY: Belongs to the nitroreductase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M63808; AA62801.1; -
 DR PIR; A38686; A38686.
 DR PDB; 1KQB; 03-APR-02.
 DR PDB; 1KQC; 03-APR-02.
 DR PDB; 1KOD; 03-APR-02.
 DR PDB; 1NEC; 31-MAR-00.
 DR InterPro: IPR000415; Nitroreductase.
 DR Pfam; PF00881; Nitroreductase, 1.
 KW Oxidoreductase; NAD; NADP; Flavoprotein; FMN; 3D-structure.
 FT NP_BIND 153 158 NAD OR NADP (BY SIMILARITY).
 FT

SQ SEQUENCE 217 AA; 23950 MW; C42AA3DB184D5D9B CRC64;

Alignment Scores:

Pred. No.: 4,87e-11 Length: 217

Score: 198.50 Matches: 51

Percent Similarity: 48.09% Conservative: 37

Best Local Similarity: 27.87% Mismatches: 56

Query Match: 12.73% Indels: 39

DB: 1 Gaps: 5

US-09-647-661-1 (1-900) x NFNB_ENTCL (1-217)

QY 193 GAGGCCATCTTCACAGATGTTTGACAGCCATTAGAGTTTCTAGTGAAGAATTAGAA 252

Db 9 LysArgHisSerThrLysAlaPheAspAlaSerLysLysLeuThrAlaGluGluAlaGlu 28

QY 253 GAAATCGCTGAAATCGCCAGCTTCCCAAGCTTTACACACGCGCCATGCGCATTTT 312

Db 29 LysIleLysThrLeuLeuGlnTy-SerProSerSerThrAsnSerGlnProTrpHisPhe 48

QY 313 GTGATGTTTACTAATAAGGATTTAAAAAACAATTCAGTCGCACAGC----- 360

Db 49 IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyThrTyr 68

QY 361 TACTTTAATGAAGAATGATTAAAGCGCTTCAGCGTTAATGTTGTTGCTCTTTAAGA 420

Db 69 ValPheAsnGluArgLysMetLeuAspAlaSerHisValValPheCysAla----- 86

QY 421 CCTAGCGAGTTGTTTACCACACGGCCATTACATGCMAAACCTTTACCGCGAGTCTTATAAG 480

Db 27 -----LysThrAlaMetAspAlaTrpLeu 95

QY 481 GTTAGAGTGATC-----CCTTCTTTT 501

Db 96 GluArgValValAspGlnGluGluAlaAspGlyArgPheAsnThrProGluAlaLysAla 115

QY 502 GCTCAAAATGCTTGGCGTGAGATTCAACACAGCATGCAGAGATTA----- 546

Db 116 AlaAsnHisLysGlyArgThrTy-PheAlaAspMetHisArgValAspLeuLysAspAsp 135

QY 547 GAAAGCTATATTAGACATGCTATATCGTGGGCGCAAAATTTGATGGCGGTGAGC 606

Db 136 AspGlnTrpMetAlaLysGlnValTyLeuAsnValGlyAsnPheLeuLeuGlyValGly 155

QY 607 TTAATGGATTGGATGCTGCTATTGAGGCTTTGATCCTTTAAAGTGGGTGAATT 666

Db 156 AlaMetGlyLeuAspAlaValProLleGluGlyPheAsp-----AlaAlaIle 171

QY 667 TTAGAAGAG 675

Db 172 LeuAspGlu 174

RESULT 6

NFNB_SALTY

ID NFNB_SALTY STANDARD; PRT; 217 AA.

AC P1588;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE OxyGen-insensitive NAD(P)H nitroreductase (EC 1.-.-.-)

GN NFNB OR STW0578.

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

NCBI_TaxID=602;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=TA1538;

RC MEDLINE=90192100; PubMed=2179862;

RA Watanabe M., Ishida M., Nozaki T.;

RT "Nucleotide sequence of Salmonella typhimurium nitroreductase gene.";

RL Nucleic Acids Res. 18:1059-1059(1990).

RN [2]

SEQUENCE FROM N.A.

RC STRAIN=LT2 / SGSC1412 / ATCC 700720;

RC MEDLINE=21534948; PubMed=11677609;

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,

RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,

RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,

RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,

RA Waterston R., Wilson R.K.;

RT "Complete genome sequence of Salmonella enterica serovar Typhimurium

RT LT2.";

RL Nature 413:852-856(2001).

CC -!- FUNCTION: REDUCTION OF A VARIETY OF NITROAROMATIC COMPOUNDS USING

CC NADH (AND TO LESSER EXTENT NADPH) AS SOURCE OF REDUCING

CC EQUIVALENTS; TWO ELECTRONS ARE TRANSFERRED. CAPABLE OF REDUCING

CC NITROFURAZONE (BY SIMILARITY).

CC -!- COFACTOR: FMN.

CC -!- SUBUNIT: Monomer (By similarity).

CC -!- SIMILARITY: Belongs to the nitroreductase family.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; X17250; CAA35113.1; -.

DR EMBL; AE008722; AAL19529.1; -.

DR PIR; S08397; S08397.

DR HSP; P38489; IDS7.

DR StyGene; SG10246; nfnB.

DR InterPro; IPR000415; Nitroreductase.

DR Pfam; PF00881; Nitroreductase; 1.

DR Oxidoreductase; NAD; NADP; Flavoprotein; FMN; Complete proteome.

FT NP_BIND 153 158 NAD OR NADP (BY SIMILARITY).

SQ SEQUENCE 217 AA; 23955 MW; E90E9E05A0826D0F CRC64;

Alignment Scores:

Pred. No.: 7,55e-11 Length: 217

Score: 196.50 Matches: 52

Percent Similarity: 48.35% Conservative: 36

Best Local Similarity: 28.57% Mismatches: 55

Query Match: 12.60% Indels: 39

DB: 1 Gaps: 5

US-09-647-661-1 (1-900) x NFNB_SALTY (1-217)

QY 193 GAGGCCATCTTCACAGATGTTTGACAGCCATTAGAGTTTCTAGTGAAGAATTAGAA 252

Db 9 GlnArgTy-SerThrLysAlaPheAspProSerLysLysLeuThrAlaGluAlaAsp 28

QY 253 GAAATCGCTGAAATCGCCAGCTTCCCAAGCTTTACACACGCGCCATGCGCATTTT 312

Db 29 LysIleLysThrLeuLeuGlnTy-SerProSerSerThrAsnSerGlnProTrpHisPhe 48

QY 313 GTGATGTTTACTAATAAGGATTTAAAAAACAATTCAGTCGCACAGC----- 360

Db 49 IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68

QY 361 TACTTTAATGAAGAATGATTAAAGCGCTTCAGCGTTAATGTTGTTGCTCTTTAAGA 420

Db 69 ThrPheAsnGluArgLysMetLeuAspAlaSerHisValValPheCysAla----- 86

QY 421 CCTAGCGAGTTGTTTACCACACGGCCATTACATGCMAAACCTTTACCGCGAGTCTTATAAG 480

Db 87 -----LysThrAlaMetAspAlaTrpLeu 95

QY 481 GTTAGAGTGATC-----CCTTCTTTT 501

Db 96 GluArgValValAspGlnGluAlaAspGlyArgPheAlaThrProGluAlaLysAla 115

QY 502 GCTCAAAATGCTTGGCGTGAGATTCAACACAGCATGCAGAGATTA----- 546

Db 116 AlaAsnAspLysGlyArgArgPheAlaAspMethHisArgValSerLeuLysAspAsp 135
 Qy 547 GAAAGCTATATTGAGCAATGCTATATCGCTGTGGGCAAAATTTGCATGGCGTGAGC 606
 Db 136 HisGlnTrpMetAlaLysGlnValTyrLeuAsnValGlyAsnPheLeuLeuGlyValAla 155
 Qy 607 TTAATGGATGGATGCTGATCATTTGAGGCTTTCATCTTTAAAGTGGGTGAAGTT 666
 Db 156 AlaMetGlyLeuAspAlaValProIleGluGlyPheAsp-----AlaGluVal 171
 Qy 667 TTAGAA 672
 Db 172 LeuAsp 173

RESULT 7
 NFNB_ECOLI STANDARD; PRT; 217 AA.
 ID PFNB_ECOLI
 AC P38489; F19575;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Oxygen-insensitive NAD(P)H nitroreductase (EC 1.-.-.-) (FMN-dependent
 DE nitroreductase) (Dihydropteridine reductase) (EC 1.6.99.7).
 GN NFNB OR NFBI OR NFSE OR NNR OR DPRA OR B0578.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / CS60;
 RA Zeno S., Koike H., Tanokura M., Saigo K.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B;
 RX MEDLINE=95113294; PubMed=7813889;
 RA Michael N.P., Brehm J.K., Anlezark G.M., Minton N.P.;
 RT "Physical characterisation of the Escherichia coli B gene encoding
 RT nitroreductase and its over-expression in Escherichia coli K12";
 RL FEMS Microbiol. Lett. 124:195-202(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horuchi T.;
 RT "A 718-Kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map";
 RL DNA Res. 3:137-155(1996).
 RN [6]
 RP CHARACTERIZATION, AND SEQUENCE OF 1-31 AND 139-180.

RC STRAIN=B;
 RX MEDLINE=931112106; PubMed=1472094;
 RA Anlezark G.M., Melton R.G., Sherwood R.F., Coles B., Friedlos F.,
 RA Knox R.J.;
 RT "The bioactivation of 5-(aziridin-1-yl)-2,4-dinitrobenzamide
 RT (CB1954) -- I. Purification and properties of a nitroreductase enzyme
 RT from Escherichia coli -- a potential enzyme for antibody-directed
 RT enzyme prodrug therapy (ADAPT)";
 RL Biochem. Pharmacol. 44:2289-2295(1992).
 RN [7]
 RP SEQUENCE OF 1-12.
 RC STRAIN=K12 / EMG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 RT in the genome of Escherichia coli K-12";
 RL Electrophoresis 18:1259-1313(1997).
 RN [8]
 RP SEQUENCE OF 1-20, AND CHARACTERIZATION AS A DIHYDROPTERIDINE REDUCT.
 RX MEDLINE=89076221; PubMed=3060113;
 RA Vasudevan S.G., Shaw D.C., Armarego W.L.F.;
 RT "Dihydropteridine reductase from Escherichia coli";
 RL Biochem. J. 255:581-588(1988).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.06 ANGSTROMS).
 RC STRAIN=B;
 RX MEDLINE=20476666; PubMed=11020276;
 RA Parkinson G.N., Skelly J.V., Neidle S.;
 RT "Crystal structure of FMN-dependent nitroreductase from Escherichia
 RT coli B: a prodrug-activating enzyme";
 RL J. Med. Chem. 43:3624-3631(2000).
 CC -1- FUNCTION: REDUCTION OF A VARIETY OF NITROAROMATIC COMPOUNDS USING
 CC NADH (AND TO LESSER EXTENT NADPH) AS SOURCE OF REDUCING
 CC EQUIVALENTS; TWO ELECTRONS ARE TRANSFERRED. CAPABLE OF
 CC NITROFUZONONE, QUINONES AND THE ANTI-TUMOR AGENT CB1954 (5-
 CC (AZIRIDIN-1-YL)-2,4-DINITROBENZAMIDE). THE REDUCTION OF CB1954
 CC RESULTS IN THE GENERATION OF CYTOTOXIC SPECIES.
 CC -1- CATALYTIC ACTIVITY: NAD(P)H + 6,7-dihydropteridine = NAD(P) (+) +
 CC 5,6,7,8-tetrahydropteridine.
 CC -1- COFACTOR: FMN.
 CC -1- SUBUNIT: Monomer or homodimer.
 CC -1- SIMILARITY: Belongs to the nitroreductase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC
 CC EMBL; D25414; BAA05004.1; -
 CC EMBL; U07860; AAC43263.1; -
 CC EMBL; AE000163; AAC73679.1; -
 CC EMBL; U92598; AAB40776.1; -
 CC EMBL; D90700; BAA35218.1; -
 CC EMBL; A23284; CAA01666.1; -
 CC EIR; I67685; I67685.
 CC FIR; S01818; S01818.
 CC FDB; I057; 12-JUL-00.
 CC FDB; I1CR; 23-MAY-01.
 CC FDB; I1CV; 23-MAY-01.
 CC FDB; I1CV; 23-MAY-01.
 CC FDB; I1CV; 23-MAY-01.
 CC EcoGene; EG20151; nfnB.
 CC InterPro; IPR000415; Nitroreductase.
 CC Pfam; PF00881; Nitroreductase; 1.
 CC Oxidoreductase; Nitroreductase; 1.
 CC 3D-structure.
 CC NP_BIND 153 158 NAD OR NADP (BY SIMILARITY).
 CC S -> C (IN REF. 8).
 CC RHS -> CIV (IN REF. 8).
 CC S -> M (IN REF. 8).
 CC
 CC FT NP_BIND 153 158
 CC FT CONFLICT 5 5
 CC FT CONFLICT 10 12
 CC FT CONFLICT 19 19

175 MetLeuIleSeIleGlyNysAlaValLysProAlaHisGlnSerAsnArgLeuProLeu 194

760 GTTGATCGCATTAAGTGGTTG 780
|||||
195 SerLysValSerThrTrpLeu 201
|||||

RESULT 9

YDGI_BACSU STANDARD; PRT; 209 AA.

ID YDGI_BACSU STANDARD; PRT; 209 AA.
P96707;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Putative NAD(P)H nitroreductase ydgi (EC 1.-.-.-).
YDGI OR BSU05660.
Bacillus subtilis.
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxID=1423;
[1]
SEQUENCE FROM N.A.
STRAIN=168;
RCA
Kasahara Y., Nakai S., Lee S., Sadaie Y., Ogasawara N.;
"A 148 kbp sequence of the region between 35 and 47 degree of the
Bacillus subtilis genome.";
Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
[2]
SEQUENCE FROM N.A.
STRAIN=168;
MEDLINE=98044033; PubMed=9384377;
Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
Azevedo V., Barbero M.G., Bessieres P., Bolotin A., Borchert S.,
Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
Denizot F., Devigne K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
Hilbert H., Holtsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
Joris B., Karamanolis D., Kasahara Y., Klaerr-Bianchard M., Klein C.,
Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V.,
Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigic C.,
Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
Parro V., Pohi T.M., Portetelle D., Porwollik S., Prescott A.M.,
Prescan E., Pujic P., Furnelle B., Rapoport G., Rey M., Reynolds S.,
Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
Sato T., Scanlan E., Schleich S., Schroter R., Scoffone F.,
Sekiguchi J., Sekowska A., Seror S.J., Serroir P., Shin B.S., Solido B.,
Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,
Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tononi A.,
Tosato V., Uchiyama S., Vandenbol M., Vannier P., Vassarotti A.,
Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
"The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis";
Nature 390:249-256(1997).
- l- COFACTOR: FMN (Potential).
- i- SIMILARITY: Belongs to the nitroreductase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to licens@isb-sib.ch).

CC
EMBL; AB001486; BAA19399.1; -
DR
YDGI; Z99107; CAB12385.1; -
CC

DR PIR: C69783; C69783.
 DR HSPF: Q60049; INOX.
 DR Subtilist; BG12176; ydgi.
 DR InterPro: IPR000415; Nitroreductase.
 DR Pfam: PF00881; Nitroreductase; 1.
 KW Hypothetical protein; Oxidoreductase; NAD; NADP; Flavoprotein; FMN;
 KW Complete proteome.
 SQ SEQUENCE 209 AA; 23359 MW; A40E9D459F560597 CRC64;
 Alignment Scores:
 Pred. No.: 1,47e-08 Length: 209
 Score: 172.50 Matches: 52
 Percent Similarity: 49.07% Conservative: 53
 Best Local Similarity: 24.30% Mismatches: 80
 Query Match: 11.06% Indels: 29
 DB: 1 Gaps: 9

US-09-647-661-1 (1-900) x YDGI_BACSU (1-209)

Qy 181 CAACATATTAAACGAGCCCATCTTCGCAAGATGTTTGACAGCCATTATGAGTTTCTAGT 240
 Db 9 GluileMetLysGlyArgArgSerIleArgAsnTyrAspProAlaValLysIleSerLys 28

Qy 241 GAAGAATTAGAAATCGCTGAATCGCAGGCTATCGCAGAGCTCTTACAACAGCAG 300
 Db 29 GluGluMetThrGluileLeuGluGluAlaThrThrAlaProSerSerValAlaGln 48

Qy 301 CCATGGCATTTGTGATGGTTACTAATAAGGATTTAAACAAACAAATTCGAGTCACAGC 360
 Db 49 ProTprArgPheLeuValIleAspSerProGluGlyLysGluLysLeuAlaProLeuAla 68

Qy 361 TACTTTATGAAGAAATGATTAAGCGCTTCAGCGTTAATGGTGGTATGCTCTTTAAGA 420
 Db 69 SerPheAsnGlnThrGlnValThrThrSerSerAlaValIleAlaValPheAla---- 86

Qy 421 CCTAGCGATGTTACACACGCGCCATTACATGACAAACCTTTAC----- 465

Db 87 -----AspMetAsnAsnAlaAspTyrLeuGluGluIleTyrSerLysAlaValGlu 103

Qy 466 -----CGGAGTCTTATPAGGTTAGATGATCCCTCTCTTTTGT----- 504

Db 104 LeuGlyTyrMetProGlnGluValLysAspArgGlnIleAlaLeuThrAlaHisPhe 123

Qy 505 CAATGCTTGGCGTGAGATTCACACACAGCATGCAAGATAGAAAGCTATATTAGAG 564
 Db 124 GluLysLeuProAlaGlnValAsnArg-----GluThrIleLeuIleAsp 138

Qy 565 CAATGCTATATCGCTGTGGGCAATTTGTCATGGCGCTTAATGGGATTCGATAGT 624
 Db 139 GlyGlyLeuValSerMet---GlnLeuMetLeuThrAlaArgAlaHisGlyTyrAspThr 157

Qy 625 TGCATTATTGAGCGCTTGATCCCTTTAAAGTGGGTGAAGTT-----TTAGAGAGCGT 678
 Db 158 AsnProIleGlyGlyTyrAspLysGluAsnIleAlaGluThrPheGlyLeuAspLysGlu 177

Qy 679 ATCAATAAGCCTAAATCGCATGCTTCATCGCTTTGGCAAGAGGGTGGCAGAGCAGC 738
 Db 178 ArgTyrValPro-----ValMetLeuLeuSerIleGlyLys---AlaAlaAspGluGly 194

Qy 739 CAAAAATCAAGAAATCAAAAGTTGATCGGATTACT---TGG 777
 Db 195 TyrAlaSerTyrArgLeuProIleAspThrIleAlaGluTrp 208

RESULT 10
 DRGA_SNNY3
 ID DRGA_SNNY3 STANDARD; PRT; 210 AA.
 AC Q55233; P73850;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Protein drga.
 GN DRGA OR SLE1719.
 OS Synechocystis sp. (strain PCC 6803).

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98316704; PubMed=9654132;
 RA Blanskava I.V., Chesnavichene E.A., Vernotte C., Astier C.;
 RT "Resistance to nitrophenolic herbicides and metronidazole in the
 cyanobacterium Synechocystis sp. PCC 6803 as a result of the
 inactivation of a nitroreductase-like protein encoded by drga gene.";
 RL FEBS Lett. 428:188-192(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
 Hosouchi I., Matsuno A., Muraki A., Nakazaki N., Nardo K.,
 Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
 Yamada M., Yasuda M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 RN [3]
 RP SEQUENCE OF 1-17.
 RX MEDLINE=97443974; PubMed=9298645;
 RA Sazuka T., Ohara O.;
 RT "Towards a proteome project of cyanobacterium Synechocystis sp.
 strain PCC6803: linking 130 protein spots with their respective
 genes.";
 RL Electrophoresis 18:1252-1258(1997).
 CC -!- FUNCTION: CONTROLS RESISTANCE TO THE HERBICIDE DINOSEB AND
 METRONIDAZOLE. INVOLVED IN DETOXIFICATION OF DINOSEB VIA THE
 REDUCTION OF THE NITRO GROUP(S) AND THIS PROCESS IS ACCOMPANIED BY
 THE FORMATION OF TOXIC SUPEROXIDE ANIONS.
 CC -!- COFACTOR: FMN (Potential).
 CC -!- SIMILARITY: Belongs to the nitroreductase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@sib-sib.ch).
 CC EMBL; L29426; AAA91952.1; -.
 DR EMBL; D90910; BAA17905.1; -.
 DR InterPro: IPR000415; Nitroreductase.
 DR Pfam: PF00881; Nitroreductase; 1.
 KW Herbicide resistance; Oxidoreductase; NAD; NADP; Flavoprotein; FMN;
 KW Complete proteome.
 FT CONFLICT 31 31 H -> L (IN REF. 2).
 FT CONFLICT 45 45 L -> H (IN REF. 2).
 FT CONFLICT 183 210 FPKRSNSFGRIPLKLLCTKVKWCLAI -> RAKGGQTPL
 BELVWNSF (IN REF. 2).
 SQ SEQUENCE 210 AA; 23703 MW; C29A1293996EBB95 CRC64;
 Alignment Scores:
 Pred. No.: 2,83e-08 Length: 210
 Score: 169.50 Matches: 53
 Percent Similarity: 46.11% Conservative: 36
 Best Local Similarity: 27.46% Mismatches: 87
 Query Match: 10.87% Indels: 17
 DB: 1 Gaps: 5

US-09-647-661-1 (1-900) x DRGA_SNNY3 (1-210)

Qy 193 GAGCGCCATCTTCGCAAGATGTTTGACAGCCATTATGAGTTTCTAGTCAAGATTAGAA 252
 Db 9 GlnArgArgSerValLysHisPheAspProAspHisArgLeuThrAlaGluGluArg 28

Qy 253 GAAATCGCTGAATCGCAGGCTATCGCAGCTCTTACACACGACCCATGTCATT 312
 Db 253 GAAATCGCTGAATCGCAGGCTATCGCAGCTCTTACACACGACCCATGTCATT 312


```

DB: 1 Gaps: 7
US-09-647-661-1 (1-900) x FRA1_VIBF1 (1-217)
Qy 187 TTAACAGAGCCGCTTCTGCAAGATGTTGACAGCCATTATGAGTTTCTAGTGAAGAA 246
Dy 8 LeuGluasnArgTyrThrSerLysLysTyrAspProSerLysLysValSerGlnGluAsp 27
Qy 247 TTAAGAAGATCGCTGAAATCGCCAGCGCTATCGCCAGCTCTTACACAGCCAGCCATGG 306
Dy 28 LeuAlaValLeuLeuGluAlaLeuArgLeuSerAlaSerSerLeuAsnSerGlnProTyr 47
Qy 307 CATTTTGTGATGTTACTTAAGATTTAAACAAATTCGAGTGCACAGCTAC--- 363
Dy 48 LysPheIleValIle---GluSerAspAlaAlaLysGlnArgMetHisAspSerPheAla 66
Qy 364 -----TTTAATGAAGAATGATTAAGCGCTTCAGCGTTAAAGTGGTATGTC 411
Dy 67 AsnMetHisGlnPheAsnGlnProHisIleLysAlaCysSerHisValIleLeuPheAla 86
Qy 412 TCTTTAAGACCTAGCGAGTTGTTACACACAGCCATTACATGCAAAACCTTTTAC--- 465
Dy 87 AsnLysLeu-----SerTyrThrArgAspAspTyrAspVal 98
Qy 466 -----CCGAGCTTATATAGCTTAGTGATC 492
Dy 99 ValLeuSerLysAlaValAlaAspLysArgIleThrGluGluGlnLysGluAlaAlaPhe 118
Qy 493 CCTTCTTTTGTCTCAATCTTGGCGTGAGATTCAACACACAGCTCAAGATTAGAAGC 552
Dy 119 AlaSerPhe---LysPheValGluLeuAsnCysAspGluAsnGlyGluHis---LysAla 136
Qy 553 TATATTTAGCAATGCTATATCGCTGTGGGCAATTTCCATGGCGGTGAGCTTAATG 612
Dy 137 TrpThrLysProGlnAlaTyrLeuAlaLeuGlyAsnAlaLeuHisThrLeuAlaArgLeu 156
Qy 613 GGATTTGATGATGTCATTTATGAGGCTTTGATCTTTAAAGTGGGTGAAGTTTAGAA 672
Dy 157 AsnIleAspSerThrThrMetGluGlyLeuAspProGluLeuLeuSerGluIlePheAla 176
Qy 673 GAGCGTATCAATAAGCCATAATCGCATGCTTGATCGCTTTGGGC----- 717
Dy 177 AspGluLeuLysGlyTyrGluCysHisValAlaLeuAlaIleGlyTyrHisHisProSer 196
Qy 718 AAGAGGTTGGCAGAGCGAGCCAAATAATCAAGAAATCAAAAGTTGATCGCATTTAC 777
Dy 197 GluAspTyrAsnAlaSerLeuProLysSerArgLysAlaPheGluAspValIleThrIle 216
Qy 778 TTG 780
Dy 217 Leu 217

RESULT 12
YM67_ARCFU
ID YM67_ARCFU STANDARD; PRT; 174 AA.
AC O28017;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative NADH dehydrogenase/NAD(P)H nitroreductase AF2267
DE (EC 1.1.1.17).
GN AF2267.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,

```

```

RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.J., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Cocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- COFACTOR: FMN (Potential).
CC -!- SIMILARITY: Belongs to the nitroreductase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; A5000948; AAB88993.1; -.
CC FIR; C69533; C69533.
CC HSP; Q60049; INOX.
CC TIGR; AF2267;
CC InterPro; IPR000415; Nitroreductase.
CC Pfam; PF00881; Nitroreductase; 1.
CC KW Hypothetical protein; Oxidoreductase; NAD; NADP; Flavoprotein; FMN;
CC Complete proteome.
CC NP BIND 107 112 NAD OR NADP (BY SIMILARITY).
CC SEQUENCE 174 AA; 20098 MW; EA8517156D73253F CRC64;
Alignment Scores:
Pred. No.: 1.05e-06 Length: 174
Score: 153.00 Matches: 45
Percent Similarity: 44.08% Conservative: 48
Best Local Similarity: 21.33% Mismatches: 72
Query Match: 9.81% Indels: 46
DB: 5
US-09-647-661-1 (1-900) x YM67_ARCFU (1-174)
Qy 169 GAAAAAAGAGCAACTATTAAACAGCGCCCATTTCTTCAAGATGTTTACAGCCATTAT 228
Dy 2 GluGluCysLeuLysMetIleTyrThrArgAGSerIleArgValTyr---SerAspAG 20
Qy 229 GAGTTTCTTAGTGAAGATTAAGAAATCGCTGAATCGCCAGGTATCGCAAGCTCT 288
Dy 21 GlnIleSerAspGluAspIleGluLysIleLeuLysAlaAlaMetLeuAlaProSerAla 40
Qy 289 TACAACACGCGCATGTCATTTGTGATGTTTACTATAAGGATTTAAAAAACAATTT 348
Dy 41 GlyAsnGluGlnProTyrPheIleValValArgAspArgGluMetLeuLysLysMet 60
Qy 349 GCAGTGCACAGCTACTTTTAAAGAAATGATTAAAGCGCTTCAGCGTTATGGTGA 408
Dy 61 SerGluAlaPheThrPheGly---GlnMetLeuProAsnAlaSerAlaIleVal 79
Qy 409 TGCTCTTTAAGACCTAGCGAGTTGTTTACCACAGCGCCATTACATGCAAAACCTTTACCG 468
Dy 80 CysAlaAspProLysLeuSerLysTyrProTyrAspMetTyrValGlnAsp----- 96
Qy 469 GAGTCTTTAAGGTTAGAGTGATGCCCTTTCTTTTGTCCAAATGCTTGGGTGAGATTCAAC 528
Dy 96 -----
Qy 529 CACAGCATGCAAGATTGAAAGCTATATTTTAGAGCAATGCTATATCGCTGTGGGCA 588
Dy 97 -----CysSerAlaAlaThrGluAsn 103
Qy 589 ATTTGTCATGGCGTGAAGCTTAATGCGATTGGATGATGTTGATTTAGAGCTTTGATCCT 648
Dy 589 ATTTGTCATGGCGTGAAGCTTAATGCGATTGGATGATGTTGATTTAGAGCTTTGATCCT

```

Db 104 IleLeuLeuAlaAlaArgCysLeuGlyIleGlySerValTrpLeuGlyValTyProArg 123
QY 649 TTAAGAGGGGTGAAGCTTTTAAAGAGCGGTATCAATAGCTAAA-----ATCGCA 699
Db 124 GluGluArgMetLysAlaLeuArgGluLeuGlyLeuProGluAenIleValValPhe 143
QY 700 TGCTTGATCGCTTTGGG-----AAGAGGGTGGCAGAGCGAGCCAAAAA 744
Db 144 SerValValSerLeuGlyTyProLysAspGluLysAspPheTyGluAlaAspAspArg 163
QY 745 TCAGAAAAATCAAAAGTTGATCGCATTTGG 777
Db 164 PheAsnProAspArgIleHisArgGluLysTrp 174
RESULT 13
YDFN_BACSU
ID YDFN_BACSU STANDARD; PRT; 206 AA.
AC P96692;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative NAD(P)H nitroreductase ydfN (EC 1.-.-.-).
GN YDFN OR BSU05480.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RT "A 148 kbp sequence of the region between 35 and 47 degree of the
RT Bacillus subtilis genome."
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunet F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Bourlier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerth I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kashara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinou S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portetelle S., Prescott A.M.,
RA Presecan E., Fujic P., Furnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan B., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Ternstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassaretti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.P., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis".
RL Nature 390:249-256(1997).
CC -I- COFACTOR: FMN (Potential).
CC -I- SIMILARITY: Belongs to the nitroreductase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB001488; BAAL1382.1; ALT_INIT.
DR EMBL; Z99106; CAB12355.1; --.
DR PIR; D69781; D69781.
DR Subtilist; BG12161; ydfN
DR InterPro; IPR000415; Nitroreductase.
DR Fram; PF00881; Nitroreductase; 1.
DR KW Hypothetical protein; Oxidoreductase; NAD; NADP; Flavoprotein; FMN;
KW Complete proteome.
SQ SEQUENCE 206 AA; 23559 MW; 99BFE7FDBD9075FC CRC64;
Alignment Scores:
Pred. No.: 3.19e-05 Length: 206
Score: 137.50 Matches: 51
Percent Similarity: 44.88% Conservative: 41
Best Local Similarity: 24.88% Mismatches: 70
Query Match: 8.82% Indels: 43
DB: 9 Gaps: 9
US-09-647-661-1 (1-900) x YDFN_BACSU (1-206)
QY 184 CTATTAAACGAGCGCCATCTTTCGAGAGTGTTCAGCCATTAGATTTCTAGTGA 243
Db 7 LeuValAsnGluArgSerAlaSerAsnPheLeuSerGlyHisProIleThrLysGlu 26
QY 244 GAATTAGAAAGAAATCGCTGAAATGCCAGGCTATCCCAAGCTTCAACACGAGCCCA 303
Db 27 AspLeuAsnGluMetPheGluLeuValAlaLeuAlaProSerAlaPheAsnLeuGlnHis 46
QY 304 TGGCATTTGTGATGTTACTAATAGGATTTAAA-----AAACAATTCAGTG 354
Db 47 ThrLysTyValThrValLeuAspGlnAspValLysGluLysLeuLysGlnAlaAlaAsn 66
QY 355 CACAGCTACTTTAATGAAGAAATGATTAAGCGCTTCAGCGCTTAATGCTGCTCT 414
Db 67 GlyGlnTy-----LysValValSerSerSerAlaValLeuValLeuGly 82
QY 415 TTAAGACCTAGCGAGTTGTACACACGCGCCATTACATGCAA-----AACCTTACCG 468
Db 83 AspLysGlnAla-----TyrGlnGlnAlaAlaAspIleTyr----- 94
QY 469 GAGCTTATAGGTTAGAGTATCCCTTTTGTCTCAATGCTTGGCGTG----- 519
Db 95 GluGlyLeuLysVal-----LeuGlyIleLeuAsnLys 105
QY 520 ---AGATTCAACACACGAGCATGCAAGATTAGAAAGCTAT----- 555
Db 106 GlnGluTyAspHisMetValGlnAspThrValSerPheTyGluAsnArgGlyGluGln 125
QY 556 -----ATTTAGAGCAATGCTATATCGCTGTGGGCAATTTGCATGGGC 600
Db 126 PheLysArgAspGluAlaIleArgAsnAlaSerLeuSerAlaMetMetPheMetLeuSer 145
QY 601 GTGAGCTTAATGGGATGATGATGCTATTATGGAGGCTTTGATCCTTTAAAGTGGGT 660
Db 146 AlaLysGluLysGlyTrpAspThrCysProMetIleGlyPheAspAlaGluAlaValLys 165
QY 661 GAAGTTTGAAGAGCGGTATCAATAGCTTAATCGCATGCTTGTATCGCTTGGCGAAG 720
Db 166 ArgIleLeuAsnIleAsp---AspGlnPheGluValMetMetIleThrIleGlyLys 184
QY 721 AGGCTGGCAGGAAGCG 735
Db 185 GluLysThrGluSer 189
RESULT 14
NOX_THETH
ID _NOX_THETH STANDARD; PRT; 205 AA.

AC Q60049; Q53306;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NADH dehydrogenase (EC 1.6.99.3) (H2O(2) forming NADH oxidase)
 DE (NADH:oxygen oxidoreductase).
 GN NOX
 OS Thermus thermophilus.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 OC Thermus.
 OX NCBI_TaxID=274;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=HB8 / ATCC 27634;
 RC MEDLINE=92249333; PubMed=1577004;
 RA Park H.-J., Kreutzer R., Reiser C.O.A., Sprinzl M.;
 RT "Molecular cloning and nucleotide sequence of the gene encoding a
 RT H2O2-forming NADH oxidase from the extreme thermophilic Thermus
 RT thermophilus HB8 and its expression in Escherichia coli.";
 RL Eur. J. Biochem. 205:875-879(1992).
 RN [2]
 RN ERRATUM.
 RP MEDLINE=93170329; PubMed=8436145;
 RA Park H.-J., Kreutzer R., Reiser C.O.A., Sprinzl M.;
 RL Eur. J. Biochem. 211:909-909(1993).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.59 ANGSTROMS).
 RC MEDLINE=96110872; PubMed=8846223;
 RA Hecht H.J., Erdmann H., Park H.-J., Sprinzl M., Schmid R.D.;
 RT "Crystal structure of NADH oxidase from Thermus thermophilus.";
 RL Nat. Struct. Biol. 2:1109-1114(1995).
 RN [4]
 RN SEQUENCE OF 1-32, AND CHARACTERIZATION.
 RC STRAIN=HB8 / ATCC 27634;
 RC MEDLINE=92249332; PubMed=1577005;
 RA Park H.-J., Reiser C.O.A., Kondrweit S., Erdmann H., Schmid R.D.,
 RA Sprinzl M.;
 RT Purification and characterization of a NADH oxidase from the
 RT thermophile Thermus thermophilus HB8.";
 RL Eur. J. Biochem. 205:891-895(1992).
 CC -!- FUNCTION: Thermotable enzyme that exhibits highest activity at pH
 CC 5.0 and can oxidize either NADH or NADPH with a preference for
 CC NADH. Can catalyze electron transfer from NADH to various electron
 CC acceptors which include, in addition to molecular oxygen,
 CC cytochrome c, 2,6-dichlorophenolindophenol, methylene blue,
 CC ferricyanide or P-nitroblue tetrazolium.
 CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
 CC -!- COFACTOR: Binds 1 FMN per subunit.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SIMILARITY: Belongs to the nitroreductase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC
 DR EMBL; X60110; CAA42707.1; -;
 DR EMBL; S55441; AAB25458.1; -;
 DR PDB; 1NOX; 12-MAR-97.
 DR InterPro; IPR000415; Nitroreductase.
 DR Pfam; PF00881; Nitroreductase; 1.
 KW Oxidoreductase; NAD; Flavoprotein; FMN; 3D-structure.
 FT HELIX 10 16
 FT STRAND 23 23
 FT HELIX 30 40
 FT TURN 41 42
 FT HELIX 46 48
 FT STRAND 52 57
 FT HELIX 60 69
 FT TURN 70 72

FT	TURN	74	74
FT	HELIX	75	78
FT	STRAND	81	87
FT	HELIX	89	94
FT	TURN	95	95
FT	HELIX	96	98
FT	TURN	99	99
FT	TURN	102	103
FT	TURN	106	106
FT	HELIX	107	121
FT	TURN	122	122
FT	HELIX	125	149
FT	TURN	150	151
FT	STRAND	153	157
FT	HELIX	162	169
FT	TURN	170	170
FT	TURN	173	174
FT	STRAND	176	184
FT	STRAND	186	186
FT	HELIX	188	196
FT	HELIX	198	201
SQ	SEQUENCE	205 AA;	22749 MW; EE6CA2DFA85B50FE CRC64;

Alignment Scores:
 Pred. No.: 3,97e-05 Length: 205
 Score: 136.50 Matches: 46
 Percent Similarity: 44.57% Conservative: 36
 Best Local Similarity: 25.00% Mismatches: 83
 Query Match: 8.76% Indels: 19
 DB: 1 Gaps: 5

US-09-647-661-1 (1-900) x NOX_THETH (1-205)

QY	247	TTAGAAGAAATCGTGAATCCAGGCTATCGCAAGCTCTTACACACGAGCGCATGG	306
DB	33	LeuArgGluIleLeuGluAlaAlaLeuArgAlaProSerAlaTrpAsnLeuGlnProTrp	52
QY	307	CATTITGTCATGCTTACTAATAAGGATTTAAAAAACAATAATGTCAGTGCACAGCTACTTT	366
DB	53	ArgIleValValValArgAspProAlaThrLysArgAlaLeuArg---GluAlaAlaPhe	71
QY	367	AATGAAGAAATGATTAAAGCGCTTCAGGTTAATGGTGGTATGCTCTTTAAGACCTAGC	426
DB	72	GlyGlnAlaHisValGluGluAlaProValValLeuValLeuTyAlaAspLeuGluAsp	91
QY	427	GAGTTGTTACCACGCGCATTCATGCAAAACCTTTACCGG-----	468
DB	92	AlaLeu-----AlaHisLeuAspGluValIleHisProGlyValGlnGlyGluArg	108
QY	469	---GAGTCTTATAAGGTTTAGAGTGATCCCTCTTTTGTCTCAATGCTTGGCGTGATTC	525
DB	109	ArgGluAlaGlnLysGlnAlaIleGlnArgAlaPheAlaAlaMet-----	123
QY	526	ACCACAGCATGCAAGATTAGAAGCTATATTTAGACCAATGCTATATCGCTGTGGG	585
DB	124	-----GlyGlnGluAlaArgLysAlaTrpAlaSerGlyGlnSerTyrIleLeuGly	141
QY	586	CAAAATTGTCATGGCGGTGAGCTTAATGGGATTTAGTGTTCATTTATTTGGAGGCTTTGAT	645
DB	142	TyrLeuLeuLeuLeuLeuGluAlaTyArgLysGlySerValProMetLeuGlyPheAsp	161
QY	646	CTTTTAAAGTGGTGAATTTTAGAGAGCGGTATCAATAAGCTTAAATCGCATGCTTG	705
DB	162	ProGluArgValArgAlaIleLeu---GlyLeuProSerArgAlaAlaIleProAlaLeu	180
QY	706	ATCGCTTTGGCAAGAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	765
DB	181	ValAlaLeuGlyTyrProAlaGluGlyTyrProSerHisArgLeuProLeuGluArg	200
QY	766	GGGATTACTTGG 777	
DB	201	ValValLeuTrp 204	

RESULT 15

```

Y120_METTH STANDARD; PRT; 191 AA.
ID Y120_METTH O26223;
AC 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Putative NADH dehydrogenase/NAD(P)H nitroreductase [EC 1.-.-.-].
DB MTH120.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare A., Wang Y., Wierzowski J., Gibson R.,
RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delta: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155 (1997).
CC -I- COFACTOR: FMN (Potential).
CC -I- SIMILARITY: Belongs to the nitroreductase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
-----
DR ENEL; AE008081; AAB84626.1; --
DR PRK; B69027; B69027.
DR HSP; Q60049; INOX.
DR InterPro; IPR000415; Nitroreductase.
DR Pfam; PF00881; Nitroreductase; 1.
DR Hypothetical protein; Oxidoreductase; NAD; NADP; Flavoprotein; FMN;
KW Complete proteome.
FT NP_BIND 127..132 NAD OR NADP (BY SIMILARITY).
SQ SEQUENCE 191 AA; 21687 MW; FAB253F75DAD84BB CRC64;
-----
Alignment Scores:
Pred. NO.: 0.000397 Length: 191
Score: 126.00 Matches: 34
Percent Similarity: 41.07% Conservative: 35
Best Local Similarity: 20.24% Mismatches: 51
Query Match: 8.08% Indels: 48
Db: 1 Gaps: 5
-----
US-09-647-661-1 (1-900) x Y120_METTH (1-191)
Qy 181 CAACCTATTAAGCAGCGCATTCGTGCAAGATTGTTGCAGCGCATATGAGTTTTCTAGT 240
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 27 GluaIaiLeysThrAArgSerilearglystyrGlnAspArg--GluValProGlu 45
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Qy 241 GAAGAATTAGAAGAAATCGCTGAANAATCGCGAGCGCTATCGCCNAAGCTCTTACACACGAG 300
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 46 GluLeuIleAspLyIleLeuAspAlalaMetCysGlyProSerAlaValAspGluArg 65
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Qy 301 CCATGGCATTTTGTCATGGTTACTATAAGGATTTAAAAAACAAAT---GCAGTGCAC 357
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 66 ProTrpHisPheIleValValArgAsnArgGluMetLeuGluLyIleProGluValHis 85
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Qy 358 AGCTACTTTAATGAAGAATGATTAAAACGCCGTTCCAGCGTTAATGGGTATGCTCTTTA 417
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 86 ProTyr-----GluValMetValIysAspAlaproValAlaIleIleValCysCys---- 102
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

```



```

GENERAL INFORMATION:
APPLICANT: Anlezark, Gillian M.
APPLICANT: Melton, Roger
APPLICANT: Sherwood, Roger
APPLICANT: Connors, Thomas
APPLICANT: Friedlos, Frank
APPLICANT: Jarman, Michael
APPLICANT: Knox, Richard
APPLICANT: Mauger, Anthony
TITLE OF INVENTION: Bacterial Nitroreductase for the
TITLE OF INVENTION: Reduction of CB 1954 and Analogues Thereof to a Cytotoxic
TITLE OF INVENTION: Form
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSER: Spencer, Frank & Schneider
STREET: Suite 300 East, 1100 New York Avenue, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20005-3955
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/504,047
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,018
FILING DATE: 27-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Schneller, John W.
REGISTRATION NUMBER: 26,031
REFERENCE/DOCKET NUMBER: KEMFJ 0017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 414-4000
TELEFAX: (202) 414-4040
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Salmonella typhimurium
US-08-504-047-3
Alignment Scores:
Pred. No.: 2,41e-15 Length: 217
Score: 190.50 Matches: 51
Percent Similarity: 48.35% Conservative: 37
Best Local Similarity: 28.02% Mismatches: 55
Query Match: 12.22% Indels: 39
DB: 1 Gaps: 5

US-09-647-661-1 (1-900) x US-08-504-047-3 (1-217)
QY 193 GAGCGCATCTTCAAGATGTTTCAGACGCAATATGAGTTCTAGTGAAGATTAGAA 252
DB 9 GlnArgTyrSerThrLysAlaPheAspProSerLysLysLeuThrAlaGluAlaAsp 28
QY 253 GAATCGCTGAATCGCAGGCTATCCGAGCTTTACACGCGCCATCGCATTTT 312
DB 29 LysileLysThrLeuLysGlnTyrSerProSerSerThrAsnSerGlnProTrpHisPhe 48
QY 313 GTGATGGTTACTAATAGGATTTTAAAAACAAATTCGATGCGACACG----- 360
DB 49 IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68

361 TACTTTAATGAAGAAATGATTAAGAGCGCTTACGCGTTAATGTTGTTGTTGTTTAA 420
DB 69 ThrPheAsnGluArgLysMetLeuAspAlaSerHisValValPheCysAlaLysThr 88
QY 421 CCTAGCGAGTTGTACACACGCGCCATTACATGCAAAACCTTTACCGGAGTCTTAA 480
DB 89 AlaAsnAsp-----AspAlaTrpLeu 95
QY 481 GTTAGAGTATGATC-----CCTTCTTTT 501
DB 96 GluArgValValAspGlnGluAspAlaAspGlyArgPheAlaThrProGluAlaLysAla 115
QY 502 GCTCAAAATGCTTGGCGTGAATTCACACGACGATGCAAGATTA----- 546
DB 116 AlaAsnAspLysGlyArgArgPhePheAlaAspMetHisArgValSerLeuLysAspAsp 135
QY 547 GAAAGCTATATTTTAGAGCAATGCTATATCGCTGGGGCAAATTTGCGATGGCGCGAGC 606
DB 136 HisGlnTrpMetAlaLysValValTyrLeuAsnValGlyAsnPheLeuLeuGlyValAla 155
QY 607 TTAATGGGATTTGATGATGCTATTTGAGGGTTTGCATCCTTTAAAGTGGGTGAAGTT 666
DB 156 AlaMetGlyLeuAspAlaValProIleGluGlyPheAsp-----AlaGluVal 171
QY 667 TTAGAA 672
DB 172 LeuAsp 173

RESULT 4
US-09-647-855-3
; Sequence 3, Application US/09087855
; Patent No. 597065
; GENERAL INFORMATION:
; APPLICANT: Anlezark, Gillian M.
; APPLICANT: Melton, Roger
; APPLICANT: Sherwood, Roger
; APPLICANT: Connors, Thomas
; APPLICANT: Friedlos, Frank
; APPLICANT: Jarman, Michael
; APPLICANT: Knox, Richard
; APPLICANT: Mauger, Anthony
; TITLE OF INVENTION: Bacterial Nitroreductase for the
; TITLE OF INVENTION: Reduction of CB 1954 and Analogues Thereof to a Cytotoxic
; TITLE OF INVENTION: Form
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spencer, Frank & Schneider
; STREET: Suite 300 East, 1100 New York Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3955
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 08/232,018
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Schneller, John W.
; REGISTRATION NUMBER: 26,031
; REFERENCE/DOCKET NUMBER: KEMFJ 0017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 414-4000
; TELEFAX: (202) 414-4040
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Salmonella typhimurium
; US-08-504-047-3
```

; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORGANISM: Salmonella typhimurium
 ; US-09-087-855-3

Alignment Scores:
 Pred. No.: 2, 41e-15 Length: 217
 Score: 190.50 Matches: 51
 Percent Similarity: 48.35% Conservatives: 37
 Best Local Similarity: 28.02% Mismatches: 55
 Query Match: 12.22% Indels: 39
 DB: 2 Gaps: 5

US-09-647-661-1 (1-900) x US-09-087-855-3 (1-217)

```

QY 193 GAGGCCATTCTTCCAGATGTTTGACAGCCATTATGAGTTCTTCTAGTGAAGATTAGAA 252
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 9 GlnArgTyrSerThrLysAlaPheAspProSerLysLysLeuThrAlaGluAlaAsp 28
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 253 GAAATCGCTGAAATCGCCAGCTATCGCCAAAGCTTTACACACGCGCATGGCATTTT 312
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 29 LysIleLysThrLeuLeuGlnTyrSerProSerSerThrAsnSerGlnProTrpHisPhe 48
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 313 GTGATGGTCTTAATAAGGATTTAAACAAACAAATTCACATGCACGAC 360
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 49 IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 TACTTTTAATGAAGAAATGATTAAGAGCTTCAGCGTTAATGGTGTGCTCTTTAAGA 420
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 69 ThrPheAsnGluArgLysMetLeuAspAlaSerHisValValValPheCysAlaLysThr 88
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 421 CCTAGCGAGTGTGTACCACACGCGCATCATACGAAACCTTTACCGGGAGTCTTTAAG 480
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 89 AlaAsnAsp-----AspAlaTrpLeu 95
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 481 GTTAGAGTGATC-----CCTCTTTT 501
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 96 GluArgValValAspGlnGluAspAlaAspGlyArgPheAlaThrProGluAlaLysAla 115
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 502 GCTCAATGCTGGCGTGGATTCACACACGATGCAAGATTA-----546
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 116 AlaAsnAspLysGlyArgPhePheAlaAspMetHisArgValSerLeuLysAspAsp 135
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 547 GAAAGCTATATTTAGAGCAATGCTATATCGTGTGGGCAAAATTTGCATGGCGGTGAC 606
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136 HisGlnTrpMetAlaLysValValTyrLeuAsnValGlyAsnPheLeuLeuGlyValAla 155
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 607 TTAATGGGATGATGATGTCATTATTGAGGCTTTTGCATCCTTTAAAGTGGGTGAAGTT 666
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 156 AlaMetGlyLeuAspAlaValProIleGluGlyPheAsp-----AlaGluVal 171
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 567 TTAGAA 672
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 172 LeuAsp 173
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  
```

RESULT 5

US-08-232-018-2
 ; Sequence 2, Application US/08232018
 ; Patent No. 5633158
 ; GENERAL INFORMATION:
 ; APPLICANT: Anlezark, Gillian M.
 ; APPLICANT: Melton, Roger
 ; APPLICANT: Sherwood, Roger
 ; APPLICANT: Connors, Thomas
 ; APPLICANT: Friedlos, Frank
 ; APPLICANT: Jarman, Michael
 ; APPLICANT: Knox, Richard
 ; APPLICANT: Mauger, Anthony

; TITLE OF INVENTION: Bacterial Nitroreductase for the
 ; TITLE OF INVENTION: Reduction of CB 1954 and Analogues Thereof to a Cytotoxic
 ; TITLE OF INVENTION: Form
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Spencer, Frank & Schneider
 ; STREET: Suite 300 East, 1100 New York Avenue, N.W.
 ; CITY: Washington, D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005-3955
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/232,018
 ; FILING DATE: 27-JUN-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Schmeidler, John W.
 ; REGISTRATION NUMBER: 26,031
 ; REFERENCE/DOCKET NUMBER: KEMPJ 0017
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 414-4000
 ; TELEFAX: (202) 414-4040
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 217 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-232-018-2

Alignment Scores:
 Pred. No.: 4, 32e-15 Length: 217
 Score: 188.50 Matches: 49
 Percent Similarity: 47.73% Conservatives: 35
 Best Local Similarity: 27.84% Mismatches: 51
 Query Match: 12.09% Indels: 41
 DB: 1 Gaps: 5

US-09-647-661-1 (1-900) x US-08-232-018-2 (1-217)

```

QY 193 GAGGCCATTCTTCCAGATGTTTGACAGCCATTATGAGTTCTTCTAGTGAAGATTAGAA 252
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 9 LysArgHisSerThrLysAlaPheAspAlaSerLysLysLeuThrProGluAlaGlu 28
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 253 GAAATCGCTGAAATCGCCAGCTATCGCCAAAGCTTCACACACGCTTACACACGCGCATGGCATTTT 312
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 29 GlnIleLysThrLeuLeuGlnTyrSerProSerSerThrAsnSerGlnProTrpHisPhe 48
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 313 GTGATGGTCTTAATAAGGATTTAAACAAACAAATTCACATGCACGAC-----360
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 49 IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 TACTTTTAATGAAGAAATGATTAAGAGCTTCAGCGTTAATGGTGTGCTCTTTAAGA 420
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 69 ValPheAsnGluArgLysMetLeuAspAlaSerHisValValValPheCysAlaLysThr 88
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 421 CCTAGCGAGTGTGTACCACACGCGCATTCATACGAAACCTTTAC-----465
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 89 Ala-----MetAspAspValTrpLeuLysLeuValVal 99
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 466 -----CCGAGTCTTTATAAGTTAGGTAGTGATC 492
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 100 AspGlnGluAspAlaAspGlyArgPheAlaThrProGluAla-----113
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 493 CTTCTTTTGTCTCAATGCTTGGGTGAGATTCAACCCACACGATGCAAGATTAGAA---549
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 ---LysAlaAlaAsnAspLysGlyArgLysPheAlaAspMetHisArgLysAspLeu 132
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 550 -----AGCTATATTTAGACAAATGCTATATCGCTGTGGGCAAAATTTGCATG 597
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  
```

```

Db      29  GlnIleLysThrLeuLeuGlnTyrSerProSerSerThrAsnSerGlnProTrpHisPhe 48
QY      313  GTGATGGTTACTAATAAGAGATTTAAAAAACAATAATGCAGTGCACAGC----- 360
Db      49  IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68
QY      361  TACTTTTAATGAAGAATGATATAAAGCGCTTCAGCGTTAATGCTGGTATGCTCTTTAAGA 420
Db      69  ValPheAsnGluArgLysMetLeuAspAlaSerHisValValValPheCysAlaLysThr 88
QY      421  CCTAGCGAGTGTTTACCACAGCGCCATTACATCCAAACACCTTTAC----- 465
Db      89  Ala-----MetAspAspValTrpLeuLysLeuValVal 99
QY      466  -----CCGAGCTCTTATAAGTTAGTAGTGATC 492
Db      100  AspGlnGluAspAlaAspGlyArgPheAlaThrProGluAla----- 113
QY      493  CCTTCITTTTGCTCAAATGCTTGGCGTGAGATTCACACACAGCATGCAAGATTAGAA--- 549
Db      114  ---LysAlaAlaAsnAspLysGlyArgLysPhePheAlaAspMetHisArgLysAspLeu 132
QY      550  -----AGCTATATTTTAGACCAATGCTATATCGCTCTGGGGCAAATTTGCATG 597
Db      133  HisAspAspAlaGluTrpMetAlaLysGlnValTyrLeuAsnValGlyAsnPheLeuLeu 152
QY      598  GCGCTGAGCTTAATGGATGGATAGTTGCATTATGGAGCGCTTGCAT 645
Db      153  GlyValAlaAlaLeuGlyLeuAspAlaValProIleGluGlyPheAsp 168

RESULT 7
US-08-640-808-2
; Sequence 2, Application US/08640808
; Patent No. 5958682
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IMPROVEMENTS RELATING TO CANCER THERAPY
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BFO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/640,808
; FILING DATE:
; CLASSIFICATION: 514
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-640-808-2

Alignment Scores:
Pred. No.: 4.32e-15 Length: 217
Score: 188.50 Matches: 49
Percent Similarity: 47.73% Conservative: 35
Best Local Similarity: 27.84% Mismatches: 51
Query Match: 12.09% Indels: 41
DB: 2 Gaps: 5

US-09-647-661-1 (1-900) x US-08-640-808-2 (1-217)
QY      193  GAGCGCATTCCTGCAAGATGTTTGCAGCGCATATGAGTTTCTAGTGAAGATTAGAA 252
Db      9  LysArgHisSerThrLysAlaPheAspAlaSerLysLysLeuThrProGluGlnAlaGlu 28
QY      253  GAAATGCTGAAATCGCAGCGTATCGCAAGCTCTTACACGCGCAGCATGCGATTTT 312
Db      29  GlnIleLysThrLeuLeuGlnTyrSerProSerSerThrAsnSerGlnProTrpHisPhe 48

```

QY 313 GTGATGGTTACTATAAGGATTAAACAAACAAATTCAGTCACAGC----- 360
Db 49 ileValalaSerThrGluGluGlyLeuAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68
QY 361 TACTTTAATGAAGAAATGATTAAAGCGCTTCACGGTTAATGGTGGTATGCTCTTTAAGA 420
Db 69 valPheasnGluArgLysMetLeuAspAlaSerHisValValPheCysAlaLysThr 88
QY 421 CCTAGCGAGTTGTTACACACAGCGCATTCATACATGCAAAACCTTTAC----- 465
Db 89 Ala-----MetAspAspValTrpLeuLysLeuValVal 99
QY 466 -----CCGGAGTCTTATAAGTTAGTTCAGTCATC 492
Db 100 AspGlnGluAspAlaAspGlyArgPheAlaThrProGluAla----- 113
QY 493 CTTCTTTTGTCAAAATCTTGGGTGAGATTCAACACAGCATGCAAGATTAGAA--- 549
Db 114 ---LysAlaAlaAsnAspLysGlyArgLysPhePheAlaAspMetHisArgLysAspLeu 132
QY 550 -----ACGTATATTTAGACGAATGCTATATCGTGGGGGCAAAATTTGCATG 597
Db 133 HisAspAlaGluTrpMetAlaLysGlnValTrpLeuAsnValGlyAsnPheLeuLeu 152
QY 598 GCGGTGCTTAATGGGATTGGATGTCATTTATTCAGCGCTTTGAT 645
Db 153 GlyValAlaAlaLeuGlyLeuAspAlaValProIleGluGlyPheAsp 168

RESULT 8
US-09-087-855-2
; Sequence 2, Application US/09087855
; Patent No. 5977065
; GENERAL INFORMATION:
; APPLICANT: Anlezark, Gillian M.
; APPLICANT: Melton, Roger
; APPLICANT: Sherwood, Roger
; APPLICANT: Connors, Thomas
; APPLICANT: Friedlos, Frank
; APPLICANT: Jarman, Michael
; APPLICANT: Knox, Richard
; APPLICANT: Mauger, Anthony
; TITLE OF INVENTION: Bacterial Nitroreductase for the
; TITLE OF INVENTION: Reduction of CB 1954 and Analogues Thereof to a Cytotoxic
; TITLE OF INVENTION: Form
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spencer, Frank & Schneider
; STREET: Suite 300 East, 1100 New York Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3955
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/087,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/232,018
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Schellier, John W.
; REGISTRATION NUMBER: 26,031
; REFERENCE/DOCKET NUMBER: KEMPJ 0017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 414-4000
; TELEFAX: (202) 414-4040
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-087-855-2
Alignment Scores:
Pred. No.: 4,328-15 Length: 217
Score: 188.50 Matches: 49
Percent Similarity: 47.73% Conservative: 35
Best Local Similarity: 27.84% Mismatches: 51
Query Match: 12.03% Indels: 41
DB: 5 Gaps: 5
US-09-647-661-1 (1-900) x US-09-087-855-2 (1-217)
QY 193 GAGCGCCATTCTTGCAGATGTTTACAGCCATTAGAGTTTCTAGTGAAGAATTAGAA 252
Db 9 LysArgHisSerThrLysAlaPheAspAlaSerLysLysLeuThrProGluGluAlaGlu 28
QY 253 GNAATCGCTGAATCGCCAGCGCTATCGCCAAAGCTCTTACACACAGCCAGCCATGTCATTT 312
Db 29 GlnIleLysThrLeuLeuGlnTrpSerProSerSerThrAsnSerGlnProTrpHisPhe 48
QY 313 GTGATGGTTACTATAAGGATTAAACAAACAAATTCAGTCACAGC----- 360
Db 49 ileValalaSerThrGluGluGlyLeuAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68
QY 361 TACTTTAATGAAGAAATGATTAAAGCGCTTCACGGTTAATGGTGGTATGCTCTTTAAGA 420
Db 69 valPheasnGluArgLysMetLeuAspAlaSerHisValValPheCysAlaLysThr 88
QY 421 CCTAGCGAGTTGTTACACACAGCGCATTCATACATGCAAAACCTTTAC----- 465
Db 89 Ala-----MetAspAspValTrpLeuLysLeuValVal 99
QY 466 -----CCGGAGTCTTATAAGTTAGTTCAGTCATC 492
Db 100 AspGlnGluAspAlaAspGlyArgPheAlaThrProGluAla----- 113
QY 493 CTTCTTTTGTCAAAATCTTGGGTGAGATTCAACACAGCATGCAAGATTAGAA--- 549
Db 114 ---LysAlaAlaAsnAspLysGlyArgLysPhePheAlaAspMetHisArgLysAspLeu 132
QY 550 -----ACGTATATTTAGACGAATGCTATATCGTGGGGGCAAAATTTGCATG 597
Db 133 HisAspAlaGluTrpMetAlaLysGlnValTrpLeuAsnValGlyAsnPheLeuLeu 152
QY 598 GCGGTGCTTAATGGGATTGGATGTCATTTATTCAGCGCTTTGAT 645
Db 153 GlyValAlaAlaLeuGlyLeuAspAlaValProIleGluGlyPheAsp 168

RESULT 9
US-09-540-236-3762
; Sequence 3762, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATP
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3762
; LENGTH: 232
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3762
Alignment Scores:
Pred. No.: 1,666-14 Length: 232
Score: 184.00 Matches: 50
Percent Similarity: 47.87% Conservative: 51

```

Best Local Similarity: 23.70% Mismatches: 102
Query Match: 11.80% Indels: 8
DB: 4 Gaps: 3
US-09-647-661-1 (1-900) x US-09-540-236-3762 (1-232)

QY 148 TCAATGAATTTTGGATCAGGAAAAGAAAGACACTATTAACGAGCGCCATCTTGC 207
DB 13 ThrMetGlnPhePheAsnLysAspValLeuAsnAlaPheHisTyrArgValSerThr 32
QY 208 AAGATGTTGACAGCCATATGAGTTTCTAGTGAAGAATTAAGAAGAAATCGCTGAATC 267
DB 33 ArgSerTyrAspGlyAspLysLysIleProAlaGluAspPheAsnTyrIleLeuGluLeu 52
QY 268 GCAGGCTATCGCAAGCTTTACACACGACCGCATGCAATTTGATGAGTTACTAT 327
DB 53 GlyArgLeuSerProSerValGlySerGluProTyrGlnPheLeuValLeuGlnAsn 72
QY 328 AAGGATTAAAAAACAATTCAGTGACAGCTACTTTAATGAAGAAATGATTAAGGC 387
DB 73 AlaAspLeuArgGlnLysLeuLysProTyrCysTyrGlyIleProThrMetGluThrSer 92
QY 388 GTTTCAGGTTAAGTGGTATGCTCTTTAAGACCTTAGCGAGTTGTACACAC---GGC 444
DB 93 SerHisIleValAlaIleLeuAlaLysLysAsnAlaArgTyrAspThrProTyrPheGly 112
QY 445 CATTACATGCAAAAC-----CTTACCAGGAGTCTTATAGGTAGAGTGCCT 495
DB 113 GluIleMetAspArgArgGlyLeuThrGlyAspAspArgAlaLysAlaLeuGluValTyr 132
QY 496 TCTTTGCTCAATGCTGGTGGTATGCTTCAACACAGCATGCAAGATAGAAAGCTAT 555
DB 133 GlnLysPheGlnGluAspIleAlaIleLeuAspSerGluArgSerLeuTyrAspTrp 152
QY 556 ATTTAGACAACTATATCGTGGTGGGCAATTTGCATGGCGCTGCTTAATGGCA 615
DB 153 AlaSerLysGlnThrTyrIleAlaLeuAlaAsnMetMetThrGlyAlaAlaMetIleGly 172
QY 616 TTGATAGTTGCATTATTTGAGGCTTTGATCTTTAAAGTGGGTGAAGTTTGAAGAG 675
DB 173 ValAspSerCysProIleGluGlyPheAlaTyrLysGluValAsnGlnLeuLeuAlaAsp 192
QY 676 CGT-----ATCAATAGCCTAAATCGCATGCTTGATCGCTTTGGCGAAGAG 723
DB 193 GluGlyLeuPheAspProAsnGluTrpAlaValSerValMetValThrPheGlyTyrArg 212
QY 724 GTGCAGACGAGCGCAAAATCAAGAAATCA 756
DB 213 AspLysGluIleLysProLysAlaArgLysThr 223

RESULT 10
US-09-489-039A-13412
; Sequence 13412, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13412
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13412

Alignment Scores:
Pred. No.: 4,51e-14 Length: 220
Score: 180.50 Matches: 49

```

```

Percent Similarity: 47.40% Conservatives: 33
Best Local Similarity: 28.32% Mismatches: 56
Query Match: 11.58% Indels: 35
DB: 4 Gaps: 5
US-09-647-661-1 (1-900) x US-09-489-039A-13412 (1-220)

QY 193 GAGCGCATCTTCCAGATGTTTGCAGCCATATGAGTTTCTAGTGAAGATTAGAA 252
DB 12 LysArgTyrSerThrLysAlaPheAspAlaThrLysLeuThrAlaGlyGluAlaGlu 31
QY 253 GAATTCGCTGAAATCGCAGGCTATCGCAAGCTCTTACACACGACGACCATGTCATTT 312
DB 32 GlnLeuLysThrLeuLeuGlnTyrSerProSerSerThrAsnSerGlnProTyrHisPhe 51
QY 313 GTGATGGTTACTATAAGGATTTAAAAACAATTCAGTGCAGTGCACAGC----- 360
DB 52 IleValAlaSerThrAspGluGlyLysAlaArgValAlaLysAlaSerGlyThrTyr 71
QY 361 TACTTTAATGAAGAAATGATTAAAAAGCGCTTCACGCTTAATGGTGTATGCTCTTTAAGA 420
DB 72 ValPheAsnGluArgLysIleLeuAspAlaSerHisValValValPheCysAla----- 89
QY 421 CCTAGCGAGTTGTTACACACGCGCATTACATGCAAAACCTTTACCGGAGCTCTTAAAG 480
DB 90 -----LysThrAlaMetAspAlaTrpLeu 98
QY 481 GTTAGAGTATCCCTTCT-----TTTGCTCAATGCTTGGCGTGAGA 522
DB 99 GlnArgValValAspGlnGluAlaAspGlyArgPheAlaThrProAspAlaLysAla 118
QY 523 TTCAACCA-----AGCATGCAAGATTAGAA----- 549
DB 119 AlaAsnHisLysGlyArgThrPhePheAlaAspMetHisArgLysGluLeuLysAspAsp 138
QY 550 ---AGCTATATTAGACCAATGCTATATCTCTGGGCAAAATTCATCGTGGCTGAGC 606
DB 139 AspGlnTyrMetAlaLysGlnValTyrLeuAsnValGlyAsnPheLeuLeuGlyValAla 158
QY 607 TTAATGGGATGAGTATGATGTCATTATTGGAGCTTTGAT 645
DB 159 AlaMetGlyLeuAspAlaValProIleGluGlyValAsp 171

RESULT 11
US-09-328-352-6977
; Sequence 6977, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6977
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6977

Alignment Scores:
Pred. No.: 1.67e-13 Length: 218
Score: 176.00 Matches: 49
Percent Similarity: 45.92% Conservatives: 41
Best Local Similarity: 25.00% Mismatches: 86
Query Match: 11.29% Indels: 20
DB: 4 Gaps: 5
US-09-647-661-1 (1-900) x US-09-328-352-6977 (1-218)

QY 184 CTTATTAAC-----GAGCGCATCTTGCAGATGTTTGACGCCATTATGAGTTT 234

```

```

Db      4 LeuLeuAsnThrVallySerArgTyrThrThrLysAlaTyrAspProGluLysIle 23
Qy      235 TCAGTGAAGAAATAGAGAAATCGCTGAAATCGCCAGGCTATCGCAAGCTCTACAA 294
Db      24 ProGlnGluLysPheAsnLysLeuLeuGluLeuLeuArgPheThrProSerValAsn 43
Qy      295 ACGCAGCATGGCATTTTGGTGGTGTACTAATAAGGATTTAAACAAACAAATTCAGTG 354
Db      44 IleGlnProThrHisPheLeuValAlaAspAsnProThrAlaLysGluArgIleAlaLys 63
Qy      355 -----CACAGCTACTTAATGAAGAATGATTAAAGCGCTTCAGCGTTAATG 402
Db      64 AlaLeuThrGlyArgTyrAlaTyrAsnAlaProLysValLeuGluSerHisThrLeu 83
Qy      403 GTGTATGCTCTTTAAGACCTAGGAGTGTGTTCACACAGCGCATACATGCAAAACCTT 462
Db      84 ValPheCysThr-----ArgThrAspIleSerPro-----GluTyrLeuAsnGlnLeu 99
Qy      463 TACCGGAGCTTATAAGGTTAGAGTATCCCTCTTTGCTCAATGCTTGGCGTGAGA 522
Db      100 LeuGluGlnAspLeuSerGlyArgPheLysAspGluLysAlaLysLeuGlyGlnLys 119
Qy      523 -----TTCAACACACAGCATGCCAAAGATTAGAAAGCTAT 555
Db      120 AspThrArgHisGlyTyrValGluPheTyrArgAsnGluGlnLysAsnLeuPheGlyTrp 139
Qy      556 ATTTTAGAGCAATGCTATATCGCTGTGGGCAAAATTTGCATGGCGGTGAGCTTAATCGGA 615
Db      140 MetGluAsnGlnThrPheIleAlaLeuGlyGlnLeuLeuPheAlaGlyLeuGluGly 159
Qy      616 TTGATAGTTCATTATGAGGCTTTGATGCTTTTAAAGTGGTGGTGAAGTTTAGAAGAG 675
Db      160 IleAspAlaThrProMetGlyGlyPheAspGluAsnValLeuAsnGluGluPheGlyLeu 179
Qy      676 CGTATCAATAAGCCATAAATCGCATGCTTGATCGCTTTGGGCAAGAGG 723
Db      180 LysGluLysGlyLeuArgSerSerValIleValSerLeuGlyTyrArg 195

RESULT 12
US-09-107-532A-5919
; Sequence 5919, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A. Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESSES:
; ADDRESSSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007

```

```

; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5919:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...218
; SEQUENCE DESCRIPTION: SEQ ID NO: 5919:
US-09-107-532A-5919

Alignment Scores:
Pred. No.:      Length:      218
Score:          Matches:    49
Percent Similarity: 45.00%
Best Local Similarity: 22.27%
Query Match:     Mismatches: 84
DB:              Indels:    37
                Gaps:      6

US-09-647-661-1 (1-900) x US-09-107-532A-5919 (1-218)

Qy      166 CAGGAAAAAGAGAGCAACTA-----TTA 189
Db      1 GlnGlnLysArgLysArgLeuIleMetGluGlnLysAsnAspPheSerThrIleMet 20
Qy      190 AACGAGCGCCATTTGCAAGATGTTGACAGCCATTATGATGTTTCTAGTGAAGAATTA 249
Db      21 ThrGlyArgArgSerValArgValThrThrAlaProSerSerValAsnMetGlnProTrpArg 40
Qy      250 GAAGAAATCGCTGAAATCGCCAGGCTATCGCAAGCTTTTACAAACAGCCAGCCATGGCAT 309
Db      41 LeuGluMetIleLysGluAlaThrThrAlaProSerSerValAsnMetGlnProTrpArg 60
Qy      310 TTTGTGATGGTTACTAATAAGGATTTAAAAACAAATTGCGAGTCAGTCAGTCAGCTACTTAAT 369
Db      61 PheValValAlaGluSerGluGluAlaLysGluLeuLeuArgProLeuIleArgPheAsn 80
Qy      370 GAAGAAATGATTAAAGCGCTTCAGCGTTAATGGTGTA-----TGCTCT 414
Db      81 ThrArgGlnAsnAspThrSerSerAlaMetValMetIlePheGlyAspMetGlnCysTyr 100
Qy      415 TTAAGACCTAGGAGCTGTTTACCACACGGCCATTACATGCAAAACCTTTACCCGAGTCT 474
Db      101 -----GluTyrGlyGluGluIleTyrAsnGlnAla 110
Qy      475 TATAAGGTTAGAGTATGCTCCT-----TCTTTTGCTCAAAATGCTTGGC 516
Db      111 TyrGluSerGlyLysMetProLysGluValArgAspGlnGlnLeuAlaAlaIleIlePro 130
Qy      517 GTGAGATTCACACACAGCATGCAAGATTAAGACCTATATTTAGACCAATGTGTATATC 576
Db      131 TyrTyrLysSerPheSerArgGluGluMetAsnAspValValLysValAspSerSerLeu 150
Qy      577 GCTGTGGGCGCAATTTGCGATGGCGTGAGCTTAATGGGATTTGATGTTGATTTATGGA 636
Db      151 AlaAlaMetGlnPheMetLeuValAlaArgAspHisGlyTyrGluThrAsnProIleGly 170
Qy      637 GGCCTTTGATCCTTTAAAAAGTGGGTGAAGTT-----TTAGAAGAGCGGTATCAATAGCCT 690
Db      171 GlyPheGluAlaAspGlnLeuAlaGluAlaPheGlyLysAspLysAspArgTyrValPro 190
Qy      691 AAATTCGCATGCTTGTGATCGCTTTGGCAAGAGGTGGCAGAGCGCAGCAAAATCAAGA 750
Db      191 ValIle-----IleLeuSerIleGlyLysAlaValGluGluGlyTyrGluSerValArg 208

RESULT 13
US-08-412-108-4
; Sequence 4, Application US/08412108

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 20, 2004, 19:12:00 ; Search time 59 Seconds

(without alignments)
8620.093 Million cell updates/sec

Title: US-09-647-661-1

Perfect score: 1559

Sequence: 1 tgcagaattttacagagagc.....gggggtgttttcaagcggtttc 900

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DRV=xlp
-Q/cgr2_1/USFTO.spool_p/US09647661/runat_20042004_135347_2622/app_query.fasta_1.1095
-DB-A_Geneseq_29Jan04 -QFWT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09647661 @CGN 1 1 81 @runat_20042004_135347_2622 -NCPU=6
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1072	68.8	210	2	Aaw98657 H. pylori
2	1071.5	68.7	209	2	Aay43259 H. pylori
3	1067	68.4	210	3	Aay90347 H. pylori
4	258	16.5	222	4	Aag82819 S. epider
5	258	16.5	222	4	Aag82690 S. epider
6	258	16.5	226	5	Abp39622 Staphyloc
7	246	15.8	284	2	Aaw98436 H. pylori
8	234	15.0	240	6	Abm72562 Staphyloc
9	210.5	13.5	221	6	Abp78332 N. gonorr
10	204.5	13.1	190	3	Aay90339 H. influe

11	198.5	12.7	201	3	AAY90344	Aay90344 C. jejuni
12	192.5	12.3	217	2	AAW13084	Aaw13084 Escherich
13	192.5	12.3	217	2	AAW13083	Aaw13083 Escherich
14	192.5	12.3	217	6	ABP70750	Abp70750 Mutant E.
15	190.5	12.2	217	6	ABP70759	Abp70759 Mutant E.
16	189.5	12.2	217	6	ABP70754	Abp70754 Mutant E.
17	189.5	12.2	217	6	ABP70755	Abp70755 Mutant E.
18	188.5	12.1	217	2	AAR34723	Aar34723 E. coli B
19	188.5	12.1	217	2	AAR3968	Aar3968 Nitroredu
20	188.5	12.1	217	2	AAR76580	Aar76580 E. coli n
21	188.5	12.1	217	4	AAU29347	Aau29347 Novel mar
22	188.5	12.1	217	6	ABP70738	Abp70738 Wild-type
23	188.5	12.1	217	6	ABP70751	Abp70751 Mutant E.
24	188.5	12.1	614	4	ABG24803	Abg24803 Novel hum
25	186.5	12.0	217	6	ABP70752	Abp70752 Mutant E.
26	185.5	11.9	202	3	AAI90336	Aay90336 B. subtil
27	185.5	11.9	217	2	AAR95880	Aar95880 Bacterium
28	184.5	11.8	217	6	ABP70753	Abp70753 Mutant E.
29	182	11.7	221	4	AAU03119	Aau03119 Streptoco
30	182	11.7	221	5	ABP27441	Abp27441 Streptoco
31	181.5	11.6	217	6	ABP70756	Abp70756 Mutant E.
32	181.5	11.6	217	6	ABP70757	Abp70757 Mutant E.
33	181.5	11.6	217	6	ABP70758	Abp70758 Mutant E.
34	176	11.3	218	6	ADA35690	Ada35690 Acinetoba
35	175.5	11.3	248	5	ABP27440	Abp27440 Streptoco
36	172.5	11.1	209	3	AAY90335	Aay90335 B. subtil
37	172	11.0	224	5	ABP30417	Abp30417 Streptoco
38	165	10.6	180	3	AAU90341	Aay90341 Synechocy
39	160	10.3	201	6	ABU00998	Abu00998 S. pneumo
40	159	10.2	210	5	ABH43245	Abh43245 Listeria
41	153.5	9.8	218	7	ADC96292	Adc96292 E. faeciu
42	153.5	9.8	356	5	ABB54815	Abb54815 Lactococc
43	153	9.7	174	3	AAY90343	Aay90343 A. fulgid
44	151.5	9.7	202	5	ABU50992	Abu50992 Helicobac
45	150	9.6	198	3	AAU81711	Aay81711 Streptoco

ALIGNMENTS

RESULT 1

AAW98657

ID AAW98657 standard; protein; 210 AA.

XX AC AAW98657;

XX DT 31-MAR-1999 (first entry)

XX DE H. pylori GHPO 455 protein.

XX KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;

XX KW peptic ulcer disease.

XX OS Helicobacter pylori.

XX PN WO9843478-A1.

XX PD 08-OCT-1998.

XX PF 01-APR-1998; 98WO-US006371.

XX PR 01-APR-1997; 97US-00833457.

XX PR 24-JUN-1997; 97US-00881227.

XX PR 29-JUL-1997; 97US-00902615.

XX (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;

XX WPI; 1998-542293/46.

XX N-PSDB; AAX14376.

PT New isolated Helicobacter polynucleotides - used to develop products for

PT the diagnosis, prevention and treatment of Helicobacter infections and
PT gastrointestinal diseases.
XX
PS Claim 8; Page 1347-1348; 2054pp; English.
XX

CC This sequence represents a Helicobacter pylori GHPO protein of the
CC invention. The polypeptides can be used for preventing or treating
CC Helicobacter infections, and gastrointestinal diseases associated with
CC these infections, including acute, chronic, and atrophic gastritis, and
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC used for the production of antibodies. The products can also be used for
CC detection and diagnosis
XX
SQ Sequence 210 AA;

Alignment Scores:
Pred. No.: 5.4e-117 Length: 210
Score: 1072.00 Matches: 207
Percent Similarity: 99.05% Conservative: 1
Best Local Similarity: 98.57% Mismatches: 2
Query Match: 68.76% Indels: 0
DB: 2 Gaps: 0

US-09-647-661-1 (1-900) x AAW98657 (1-210)

QY 151 ATGAAATTTTGGATCGAGAAAGACACACTATTAAACGAGCGCATTTCTTGCAG 210
Db 1 MetLysPheLeuAspGlnGluLysArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTTACAGCCATTATGATTTCTAGTGAAGAAATAGAGAAATCGCTGAAATCGCG 270
Db 21 MetPheAspSerHisTyrGluPheSerSerThrGluLeuGluGluLeuAla 40
QY 271 AGGCTATCGCCAGCTCTTACACAGCAGCCATGCGATTTTGTGATGTTACTATAG 330
Db 41 ArgLeuSerProSerSerTyrAsnThrGlnProTrpHisPheValMetValThrAspLys 60
QY 331 GATTAAAAAACAATTTGCAGTGCACAGCTACTTTAAATCAAGAAATGATTAAGCGCT 390
Db 61 AspLeuLysLysGlnIleAlaHisSerTyrPheAsnGluLysSerAla 80
QY 391 TCAGCGTAATGGTGGTATGCTCTTTAAGACCTAGCAGGTTGTACACAGCGCATTTAC 450
Db 81 SerAlaLeuMetValValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTyr 100
QY 451 ATGCAAAACCTTTACCCGGAGTCTTATAAGTTAGATGATCCCTCTTTTGTCTCAAATG 510
Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
QY 511 CTGCGCTGAGATTCAACACAGCATGCAAGATTAGAAAGCTATATTTAGAGCAATGC 570
Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGlnCys 140
QY 571 TATATCCCTGTGGGCAAAATTTGATGGCGTGCAGCTTAATGGATTGGATGCTGATTC 630
Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGGAGCTTTGATCCCTTTAAAGTGGGTGAAGTTTATAGAGAGCGTATCAATAAGCCT 690
Db 161 IleGlyGlyPheAspProLeuLysValGlyValLeuGluGluArgIleAsnLysPro 180
QY 691 AAAATCCAGCTGTGATCGCTTTGGGCAAGAGGTGGCAGAGCGAGCAAAATCAAGA 750
Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
QY 751 AAATCAAAAGTGTGCGATTACTTGTGTTG 780
Db 201 LysSerLysValAspAlaIleThrTrpLeu 210

RESULT 2
AAV43259
ID AAV43259 standard; protein; 209 AA.
XX

AC AAV43259;
XX
DT 17-JAN-2000 (first entry)
XX
DE H. pylori nitroreductase, rdxA, protein sequence.
XX
KW Nitroreductase; rdxA; NADPH; nicotinamide-adenine dinucleotide phosphate;
KW electron donor; prodrug conversion; cytotoxic compound generation;
KW growth inhibitor; tumour cell; neoplastic cell.
XX
OS Helicobacter pylori.
XX
PN WO9951270-A1.
XX
PD 14-OCT-1999.
XX
PF 06-APR-1999; 99WO-US007546.
XX
PR 06-APR-1998; 98US-0080917P.
PR 14-APR-1998; 98US-0081778P.
XX
PA (UYDA-) UNIV DALHOUSIE.
XX
PI Goodwin A, Hoffman PS;
XX
DR WPI; 1999-620172/53.
DR N-PSDB; AAZ31668.
XX
PT New nitroreductase that can convert aromatic nitro prodrugs to cytotoxic
XX compounds, e.g. for killing or inhibiting cancer cells and microbes.
PS Disclosure; Fig 1; 40pp; English.

CC This sequence is the helicobacter pylori nitroreductase, designated rdxA,
CC of the invention. The nitroreductase has an isoelectric point (pI) over 6
CC ; contains at least two Cys; has a preference for NADPH (nicotinamide-
CC adenine dinucleotide phosphate) as electron donor; and can convert a
CC prodrug to one or more cytotoxic compounds. The nitroreductase is used to
CC generate cytotoxic compounds for killing or inhibiting growth of targeted
CC cells (bacteria, (retro) viruses, fungi, immune system cells, tumour
CC cells, neoplastic or other diseased cells, or cells from a particular
CC tissue or organ), especially where conjugated with a targeting compound.
CC The nitroreductase can also be used to select against cells that express
CC a functional nitroreductase-encoding gene, e.g. to detect Helicobacter
CC pylori that are resistant to metronidazole and to identify substrates for
CC nitroreductase
XX
SQ Sequence 209 AA;

Alignment Scores:
Pred. No.: 6.18e-117 Length: 209
Score: 1071.50 Matches: 209
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 0
Query Match: 68.73% Indels: 1
DB: 2 Gaps: 1

US-09-647-661-1 (1-900) x AAV43259 (1-209)

QY 151 ATGAATTTTGGATCGAGAAAGACACACTATTAAACGAGCGCATTTCTTGCAG 210
Db 1 MetLysPheLeuAspGlnGluLysArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTGACAGCCATTATGATTTTCTAGTGAAGAAATAGAGAAATCGCTGAAATCGCC 270
Db 21 MetPheAspSerHisTyrGluPheSerSerGluLeuGluIleAlaGluIleAla 40
QY 271 AGGCTATCGCCAGCTCTTACACAGCAGCCATGCGATTTTGTGATGTTACTATAAG 330
Db 41 ArgLeuSerProSerSerTyrAsnThr---ProTrpHisPheValMetValThrAsnLys 59
QY 331 GATTAAAAAACAATTTGCAAGTGCACAGCTACTTTAATGAAGAAATGATTAAGCGCT 390

OS Staphylococcus epidermidis.
XX WO200134809-A2.
PN 17-MAY-2001.
PD 09-NOV-2000; 2000WO-US030782.
PF 09-NOV-1999; 99US-0164258P.
PR (GLAX) GLAXO GROUP LTD.
PX Kimmerly WJ;
PY WPI; 2001-316495/33.
PZ N-PSDB; AAH53669.
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis.
PT Claim 18; Page 712; 2188pp; English.
PS AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce host cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH5091 to AAH5098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464
XX Sequence 222 AA;
SQ

Alignment Scores:
Pred. No.: 4,7e-21 Length: 222
Score: 258.00 Matches: 63
Percent Similarity: 53.11% Conservative: 48
Best Local Similarity: 30.14% Mismatches: 80
Query Match: 16.55% Indels: 18
DB: 4 Gaps: 6

US-09-647-661-1 (1-900) x AAG82819 (1-222)

QY 196 CGCATTTCTTGCAGATGTTTTCACAGCCATTATGATGTTTCTAGTGAAGATTAGAAGAA 255
DB 17 ArgHisAlaThrLysGluPheAspProThrLysLysIleSerAspGluAspPheAsnThr 36
QY 256 ATCGCTGAATCGCCAGGCTATGCCAGCTTTACACACCCAGCCATGGATTTGTG 315
DB 37 IleLeuGluThrGlyArgLeuSerProSerSerLeuGlyLeuGluProThrHisPheVal 56
QY 316 ATGGTTACTAATAAGATTATAAATAACAAATTCAGTCAGTCAGCTACTTTAATGAAGAA 375
DB 57 ValValaGlnAsnLysGluLeuArgGluLysLeuLysAlaTyrSerTgplAlaGlnLys 76
QY 376 ATGATTAAAGCCTTCAGGTTATGGTGGTATGCTCTTTAAGCCTAGCAGATTGTTA 435
DB 77 GlnLeuAspThrAlaSerHisPheValLeuIlePheAlaArgLys-----AsnValThr 94
QY 436 CCACAGCGGCATACATGCAAAACCTTTAC-----CCGGAGCTCTTATAAGCTTAGAGTG 489

DB 95 AlaHisThrAspTyrValGlnHisLeuLeuArgGlyValLysLysTyrGluGluSerThr 114
QY 490 ATCCCTTCT-----TTTGCTCAAAATGCTTGGCGTGAGATTCAAC 528
DB 115 IleProAlaValGluAsnLysPheAspPheGlnGluSerPheHisIleAlaAspAsn 134
QY 529 CACAGCATGCAAGATTAGAAAGCTATATTTTACCAATGCTATATCGCTGTGGGGCAA 588
DB 135 Glu-----ArgThrLeuTyrAspTgplAlaSerLysGlnThrTyrIleAlaLeuAlaAsn 152
QY 589 ATTGTCATGGCGTGAGCTTAATGGATTGATGATGATGATGATGATGATGATGATGATGAT 648
DB 153 MetMetThrSerAlaAlaLeuLeuGlyIleAspSerCysProIleGluGlyPheAspLeu 172
QY 649 TTAAGTGGTGAGCTTTAGAGAGCGT-----ATCAATAGCCCTAAATC 696
DB 173 AspLysValThrGluIleLeuSerAspGluGlyValLeuAspThrGluGlnPheGlyIle 192
QY 697 GCATGCTGTGATCGCTTTGGGCAAGAGGGTGGCAGAGCGAGCCAA---AAATCAAGAAA 753
DB 193 SerValMetValGlyPheGlyTyrArgAlaGlnGluProLysHisGlyLysValArgGln 212
QY 754 TCAAAAGTTGATCGATTACTTCTGTTG 780
DB 213 AsnGluAspAspIleIleSerTrpIle 221

RESULT 5
AAG82690
ID AAG82690 standard; protein; 222 AA.
AC AAG82690;
XX
XX 03-SEP-2001 (first entry)
DT S. epidermidis open reading frame protein sequence SEQ ID NO:2474.
XX Staphylococcus epidermidis S81 strain; infection; diagnosis; vaccination;
XX endocarditis.
OS Staphylococcus epidermidis.
XX WO200134809-A2.
XX 17-MAY-2001.
XX 09-NOV-2000; 2000WO-US030782.
XX 09-NOV-1999; 99US-0164258P.
PA (GLAX) GLAXO GROUP LTD.
XX Kimmerly WJ;
XX WPI; 2001-316495/33.
DR N-PSDB; AAH53540.
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis.
XX Claim 18; Page 652; 2188pp; English.
PS AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce host cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA

CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
 CC represent oligonucleotide sequences and primers which are used in the
 CC exemplification of the present invention. N.B. The present invention
 CC specifically claims all the polynucleotide sequences given in the
 CC sequence listing of the present specification, however the sequence
 CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
 CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
 CC for SEQ ID NO:4455 to 4464
 XX
 XX
 SQ Sequence 222 AA;

Alignment Scores:

Pred. No.: 4,73e-21 Length: 222
 Score: 258.00 Matches: 63
 Percent Similarity: 53.11% Conservative: 48
 Best Local Similarity: 30.14% Mismatches: 80
 Query Match: 16.55% Indels: 18
 DB: Gaps: 6

US-09-647-661-1 (1-900) x AAG82690 (1-222)

QY 196 CGCATTTCTTGCAGATGTTTACAGCCATTATGAGTTTCTAGTGAAGAATTAGAAGAA 255
 Db 17 ArgHisAlaThrLysGluPheAspProThrLysLysIleSerAspGluAspPheAsnThr 36
 QY 256 ATCGCTGAATGCCAGGCTATCGCAAGCTTTCACACAGCAGCCATGCGCATTTGTG 315
 Db 37 IleLeuGluThrGlyArgLeuSerProSerSerLeuGlyLeuGluProThrHisPheVal 56
 QY 316 ATGTTTACTAATAAGCATTTAAACAAATTCAGTCACAGCTTACTTAAATGAAGAA 375
 Db 57 ValValGlnAsnLysGluLeuArgGlyLysLeuAlaTyrSerTrpGlyAlaGlnLys 76
 QY 376 ATGATTAAAGCGCTTCAGCGTTAATGGTGGTATGCTTTAAGACCTAGCGAGTTGTTA 435
 Db 77 GlnLeuAspThrAlaSerHisPheValLeuPheAlaArgLys-----AsnValThr 94
 QY 436 CCACAGCGCCATTACATGCAAAACCTTTAC-----CCGAGCTCTTATAAGTTAGAGTG 489
 Db 95 AlaHisThrAspTyrValGlnHisLeuLeuArgGlyValLysLysTyrGluGluSerThr 114
 QY 490 ATCCCTTCT-----TTTGCTCAAAATGCTTGGCGTGAGATTCAC 528
 Db 115 IleProAlaValGluAsnLysPheAspPheGlnGluSerPheHisIleAlaAspAsn 134
 QY 529 CACAGCATGCAAGATTAGAACCTATATTTAGACCAATGCTATATCCTGTGGGCAA 588
 Db 135 Glu-----ArgThrLeuTyrAspTrpAlaSerLysGlnThrTyrIleAlaLeuAlaAsn 152
 QY 589 ATTTCATGCGCTGAGCTTAATGGGATTGGATGATGATTCATTTATGCGGCTTGATCCT 648
 Db 153 MetMetThrSerAlaAlaLeuLeuGlyIleAspSerCysProIleGluGlyPheAspLeu 172
 QY 649 TTAAGAGTGGTGAAGTTTATAGAGAGCGT-----ATCAATAGCCTAAATC 696
 Db 173 AspLysValThrGluIleLeuSerAspGluGlyValLeuAspThrGluGlnPheGlyIle 192
 QY 697 GCATGCTTGATCGCTTGGGCAAGAGGTGGCAGAGCGAGCCAA--AAATCAAGAAA 753
 Db 193 SerValMetValGlyPheGlyTyrArgAlaGlnGluProLysHisGlyLysValArgGln 712
 QY 754 TCAAAAGTTCATGCGATTACTTGTGTT 780
 Db 213 AsnGluAspAspIleIleSerTrpIle 221
 RESULT 6
 ABP39622
 ID ABP39622 standard; protein; 226 AA.
 XX
 AC ABP39622;
 XX
 DT 24-JUL-2002 (first entry)
 XX

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4467.
 XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.
 KW Staphylococcus epidermidis.
 OS Staphylococcus epidermidis.
 XX
 PN US6380370-B1.
 XX
 XX 30-APR-2002.
 XX
 XX 13-AUG-1998; 98US-00134001.
 PF
 XX 14-AUG-1997; 97US-0055779P.
 PR 08-NOV-1997; 97US-0064964P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX
 XX Doucette-Stamm LA, Bush D;
 PI
 XX WPI; 2002-381255/41.
 DR N-PSDB; ABN92167.
 DR
 XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
 PT polypeptide, useful for diagnosing and treating bacterial infections.
 XX
 PS Disclosure; SEQ ID NO 4467; 267pp; English.
 XX
 CC ABN90338 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences can
 CC also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to
 CC screen for compounds able to interfere with the S. epidermidis life cycle
 CC or inhibit S. epidermidis infection. N.B. The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site
 XX
 SQ Sequence 226 AA;

Alignment Scores:

Pred. No.: 4,73e-21 Length: 226
 Score: 258.00 Matches: 63
 Percent Similarity: 53.11% Conservative: 48
 Best Local Similarity: 30.14% Mismatches: 80
 Query Match: 16.55% Indels: 18
 DB: Gaps: 6

US-09-647-661-1 (1-900) x ABP39622 (1-226)

QY 196 CGCATTTCTTGCAGATGTTTACAGCCATTATGAGTTTCTAGTGAAGAATTAGAAGAA 255
 Db 21 ArgHisAlaThrLysGluPheAspProThrLysLysIleSerAspGluAspPheAsnThr 40
 QY 256 ATCGCTGAATGCCAGGCTATCGCAAGCTTTCACACAGCAGCCATGCGCATTTGTG 315
 Db 41 IleLeuGluThrGlyArgLeuSerProSerSerLeuGlyLeuGluProThrHisPheVal 60
 QY 316 ATGTTTACTAATAAGCATTTAAACAAATTCAGTCACAGCTTACTTAAATGAAGAA 375
 Db 61 ValValGlnAsnLysGluLeuArgGlyLysLeuAlaTyrSerTrpGlyAlaGlnLys 80
 QY 376 ATGATTAAAGCGCTTTCAGCGTTAATGGTGGTATGCTTTAAGACCTAGCGAGTTGTTA 435
 Db 81 GlnLeuAspThrAlaSerHisPheValLeuPheAlaArgLys-----AsnValThr 98
 QY 436 CCACAGCGCCATTACATGCAAAACCTTTAC-----CCGAGCTCTTATAAGTTAGAGTG 489
 Db 99 AlaHisThrAspTyrValGlnHisLeuLeuArgGlyValLysLysTyrGluGluSerThr 118
 QY 490 ATCCCTTCT-----TTTGCTCAAAATGCTTGGCGTGAGATTCAC 528
 XX

Db 119 IleProAlaValGluAenLysPheAspPheGlnGluSerPheHisIleAlaAspAsn 138
QY 529 CACAGCATGCAAGATTAGAAAGCTATATTTAGAGCAATGCTATATCGTGTGGGCAA 588
Db 139 Glu-----ArgThrLeuTyrrAspTrpAlaSerLysGlnThrTyrrIleAlaLeuAlaAsn 156
QY 589 ATTTCATGCGCGTGAAGCTTAATGGGATTCGATAGTTGCATTATTCGAGGCTTTGATCCT 648
Db 157 MetMetThrSerAlaAlaLeuLeuGlyIleAspSerCysProIleGluGlyPheAspLeu 176
QY 649 TTAAGCTGGGTGAAGTTTGAAGACGGT-----ATCAATAGCCTAAATC 696
Db 177 AspLysValThrGluIleLeuSerAspGluGlyValLeuAspThrGluGlnPheGlyIle 196
QY 697 GCATGCTGTATCGCTTTGGCAAGAGGTGGCAGAGCGAGCCAA---AAATCAAGAAAA 753
Db 197 SerValMetValGlyPheGlyTyrrArgAlaGlnGluProLysHisGlyLysValArgGln 216
QY 754 TCAGAACTTGATCGGATTACTTGGTTG 780
Db 217 AsnGluAspAspIleIleSerTrpIle 225
RESULT 7
AAW98436
ID AAW98436 standard; protein; 284 AA.
XX AC AAW98436;
XX AC AAW98436;
DT 31-MAR-1999 (first entry)
XX H. pylori GHPO 456 protein.
XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW peptic ulcer disease.
XX Helicobacter pylori.
XX WO9843478-A1.
XX 08-OCT-1998.
XX 01-APR-1998; 98WO-US006371.
XX 01-APR-1997; 97US-00833457.
PR 24-JUN-1997; 97US-00881227.
PR 29-JUL-1997; 97US-00902615.
XX (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
XX WPI; 1998-542293/46.
DR N-PSDB; AAX14155.
XX New isolated Helicobacter polynucleotides - used to develop products for
PT the diagnosis, prevention and treatment of Helicobacter infections and
PT gastrointestinal diseases.
XX Claim 8; Page 743-744; 2054pp; English.
XX This sequence represents a Helicobacter pylori GHPO protein of the
CC invention. The polypeptides can be used for preventing or treating
CC Helicobacter infections, and gastroduodenal diseases associated with
CC these infections, including acute, chronic, and atrophic gastritis, and
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC used for the production of antibodies. The products can also be used for
CC detection and diagnosis
XX
SQ Sequence 284 AA;

Alignment Scores:
Pred. No.: 1.32e-19 Length: 284

Score: 246.00 Matches: 49
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.00% Mismatches: 0
Query Match: 15.78% Indels: 0
DB: 2 Gaps: 0
US-09-647-661-1 (1-900) x AAW98436 (1-284)
QY 2 GCAGAAATTTTACAGAGCCAGATAGCCAAATGGGGTTTATTTTAAATTTGAGCATG 61
Db 235 AlaGluPheTyrrArgGluProAspSerGlnMetGlyValTyrrPheLeuAsnLeuSerMet 254
QY 62 GGGCAGATTTTAAGCTTATTTATGCTAGTCTTCTTCATTAGGAGTTTATTTGATGCTACA 121
Db 255 GlyGlnIleLeuSerLeuPheMetValIleValSerLeuGlyIleLeuLeuTyrrAlaThr 274
QY 122 AAAAATTTCTAAAAATAAAGCAAAATCAA 151
Db 275 LysAsnSerLysLysIleLysGluAsnGln 284
RESULT 8
ABM72562
ID ABM72562 standard; protein; 240 AA.
XX AC ABM72562;
XX AC ABM72562;
DT 20-NOV-2003 (first entry)
XX Staphylococcus aureus protein #1802.
XX Antibacterial; vaccine; Gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target.
XX Staphylococcus aureus.
XX WO200294869-A2.
XX 28-NOV-2002.
XX 27-MAR-2002; 2002WO-IB002637.
XX 27-MAR-2001; 2001GB-00007661.
XX (CHIR-) CHIRON SPA.
XX Masignani V, Mora M, Scarselli M;
XX WPI; 2003-120786/11.
DR N-PSDB; ACF74122.
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX Claim 1; SEQ ID NO 3604; 49pp; English.
XX The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus proteins of the invention
XX
SQ Sequence 240 AA;

Alignment Scores:
Pred. No.: 3.25e-18 Length: 240
Score: 234.00 Matches: 60
Percent Similarity: 49.35% Conservative: 54


```
Best Local Similarity: 25.97% Mismatches: 99
Query Match: 15.01% Indels: 18
DB: 6 Gaps: 6

US-09-647-661-1 (1-900) x ABP72562 (1-240)

QY 136 AATAAAGGAAATCAATGAAATTTTGGATCAGGAAAAAGAGACAACTATTAAACGAG 195
   |||||
DB 8 ASnLysTyrLysAsnGlyA-gLeuCysSerMetSerAsnMetAsnGlnThrIleMetAsp 27

QY 196 -----GCCATTCTTGAAGATCTTTGACAGCCATTATGAGTTTCTAGTGAA 243
   |||||
DB 28 AlaPheHisPheArgHisAlaThrLysGlnPheAspProGlnLysLysValSerLysGlu 47

QY 244 GAATTAGAAGAAATCGCTGAAATCGCCAGGCTATCGCCAGCTTACACACGCGAGCA 303
   |||||
DB 48 AspPheGluThrIleLeuGluSerGlyArgLeuSerProSerSerLeuGluGluPro 67

QY 304 TGGCAATTTGTGATGTTACTAATAAGGATTTAAAAAACAATTCAGTGCACACTAC 363
   |||||
DB 68 TrpLysPheValValIleGlnAspGlnAlaLeuArgAspGluLeuLysAlaHisSerTrp 87

QY 364 TTTAATGAAGAATGATTAAAGCGCTTCAGCGTTAATGGTGTATGCTCTTTAAGACT 423
   |||||
DB 88 GlyAlaLysGlnLeuAspThrAlaSerHisPheValLeuIlePheAlaArgLys--- 106

QY 424 AGCGAGTTGTATACACACGCGCCATTACATGCAAAACCTTTACCCGGAG-----TCTTAT 477
   |||||
DB 107 ---AsnValThrSerArgSerProTyrValGlnHisMetLeuArgAspIleLysLysTyr 125

QY 478 AAGTTAGATGATCCCTCTTTGCTCAAACTGCTGGCGTG-----AGA 522
   |||||
DB 126 GluAlaGlnThrIleProAlaValGlnGlnLysPheAspAlaPheGlnAlaAspPheHis 145

QY 523 TTCACACACAGCATGCAAGAAATAGAAAGCTATATTATAGAGCAATGCTATATCGGTGTG 582
   |||||
DB 146 IleSerAspAsnAspGlnAlaLeuTyrAspTrpSerSerLysGlnThrTyrIleAlaLeu 165

QY 583 GGGCAATTTGCATGGCGGTGAGCTTAATGGGATTCGATAGTTGCAATTTGAGGCTTT 642
   |||||
DB 166 GlyAsnMetMetThrThrAlaAlaLeuLeuGlyIleAspSerCysProMetGluGlyPhe 185

QY 643 GATCCTTTAAAAGTGGTGAAGTTTGAAGAGCGT-----ATCAATAAGCT 690
   |||||
DB 186 SerLeuAspThrValThrAspIleLeuAlaAsnLysGlyIleLeuAspThrGluGlnPhe 205

QY 691 AAAATCGCATGTTGATCGCTTTGGGCAAGAGGTGCGCAGAA---CGGAGCCAAAATCA 747
   |||||
DB 206 GlyLeuSerValMetValAlaPheGlyTyrArgGlnGlnGluProProLysAsnLysThr 225

QY 748 AGAAATCAAAAGTTGATGCGATTACTTGGTTG 780
   |||||
DB 226 ArgGlnAlaTyrGluAspValIleGluTrpVal 236

RESULT 9
ABP78332
ID ABP78332 standard; protein; 221 AA.
XX
AC ABP78332;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae amino acid sequence SEQ ID 3194.
XX
KW Antibacterial; infection; vaccine; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
FN W0200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB002069.

XX 12-FEB-2001; 2001GB-00003424.
XX (CHIR-) CHIRON SPA.
XX Fontana MR, Pizza M, Masignani V, Monaci E;
XX WPI; 2003-058415/05.
XX DR N-PSDB; ABZ39302.
XX
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection.
XX
XX Disclosure; Page 422; 815pp; English.
XX
XX The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records ABP76736-ABP81046 represent nucleic acid
XX molecules of the invention
XX
XX Sequence 221 AA;

Alignment Scores:
Pred. No.: 1,85e-15 Length: 221
Score: 210.50 Matches: 58
Percent Similarity: 46.46% Conservative: 47
Best Local Similarity: 25.66% Mismatches: 98
Query Match: 13.50% Indels: 23
DB: 6 Gaps: 4

US-09-647-661-1 (1-900) x ABP78332 (1-221)

QY 151 ATGAAATTTTGGATTCAGGAAAAAGAGACAACTATTAAACGAGCGCCATTCTTGCAG 210
   |||||
DB 1 MetThrValLeuSerLysGluGlnValLeuSerAlaPheLysAsnArgLysSerCysArg 20

QY 211 ATGTTTACACGCCATTATGAGTTTCTAGTGAACAATTAGAAGAAATCGCTGAATCGCC 270
   |||||
DB 21 HisTyrAspAlaAlaArgLysIleSerAlaGluAspPheGlnPheIleLeuGluLeuGly 40

QY 271 AGGTATCGCCAAAGCTTTTACACACGAGCGCATTTGTGATTTGTGATTTGTGATTTGTG 330
   |||||
DB 41 ArgLeuSerProSerSerValGlySerGluProTrpGlnPheValValValGlnAsnPro 60

QY 331 GATTTAAAAAACAATTCGAGTGCACAGCTACTTTAATGAAGAAATGATTAAAGCGCT 390
   |||||
DB 61 GluIleArgGlnAlaIleLysLeuPheSerTrpGlyMetAlaAspAlaLeuAspThrAla 80

QY 391 TCAGCGTTAATGGTGTATGCTCTTTAAGACCTAGCGAGTTGTTACCACAGCGCCATTAC 450
   |||||
DB 81 SerHisLeuValValPheLeuAlaLysLys-----AsnAlaArgPhe 94

QY 451 ATGCAAAACCTTTACCCGGAGCTTATAGGTTAGA-----GTG 489
   |||||
DB 95 AspSerProPheMetLeuGluSerLeuLysArgGlyValThrGluProAspAlaVal 114

QY 490 ATCCCTTCTTTTGTCT-----CAAAATGCTTGGCGTGCAGATTCAACACAGC 534
   |||||
DB 115 GluLysSerLeuAlaArgTyrGlnAlaPheGlnAlaAspAspIleLysIleLeuAspAsp 134

QY 535 ATGCAAGATTAGAAAGCTATATTTAGACCAATGCTATATCGCTGTGGGCAATTTGC 594
   |||||
DB 135 SerArgAlaLeuPheAspTrpCysArgGlnThrTyrIleAlaLeuAlaAsnMetMet 154

QY 595 ATGGCGGTGAGCTTAATGGGATTTGGATTTGCAATTTATGGAGCTTTGATCTTTTAAA 654
   |||||
DB 155 ThrGlyAlaAlaMetAlaGlyIleAspSerCysProValGluGlyPheAsnTyrAlaAsp 174

QY 655 GTGGGTGAAGTTTAGAAGAGCGGTATCAAT-----AAGCCTAAATCGCA 699
```

```
Db 175 MetGluArgValLeuSerGlyGlnPheGlyLeuPheAspAlaAlaGluTrpGlyValSer 194
Qy 700 TGCTTGATCGTTGGCAAGAGGTGGCAGAAGCGCCAAAATCAAGAAATCAAAA 759
Db 195 ValAlaAlaThrPheGlyTyArgValGlnGlnGluLeuValThrLysAlaAlaArgProLeu 214
Qy 760 GTTGATCGCATTTGG 777
Db 215 GluGluThrValIleTrp 220

RESULT 10
ID AAY90339 standard; protein; 130 AA.
XX AAY90339;
AC AAY90339;
DT 22-NOV-2000 (first entry)
DE H. influenzae nitroreductase.
KW Nitroreductase; CB1954; cytotoxic 4-hydroxylamine; 4HX derivative;
KW monofunctional alkylating agent; prodrug therapy; tumour therapy;
KW non-cytotoxic 2-hydroxylamine derivative; enzyme.
XX Haemophilus influenzae.
OS WO200047725-A1.
PN 17-AUG-2000.
FD 10-FEB-2000; 2000WO-GB0000431.
PF 10-FEB-1999; 99GB-00003019.
PR (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA Minton N, Anlezark G, Vaughan T;
PI WPI; 2000-558214/51.
PT A nitroreductase useful for producing a medicament for anti-tumour therapy
PT preferentially reduces CB1954 to a cytotoxic 4-hydroxylamine derivative
PT instead of a non-cytotoxic 2-hydroxylamine derivative.
XX Disclosure; Page 44; 56pp; English.
XX This sequence represents a nitroreductase. The invention relates to a
CC nitroreductase characterised in that it preferentially reduces the
CC monofunctional alkylating agent CB1954 to a cytotoxic 4-hydroxylamine
CC (4HX) derivative instead of a non-cytotoxic 2-hydroxylamine derivative.
CC The DNA sequence coding for the nitroreductase can be used in the
CC manufacture of a medicament for prodrug therapy. The nitroreductase,
CC preferably conjugated to a targeting moiety, can be used in the
CC manufacture of a medicament for anti-tumour therapy. The active drug is
CC not administered, it will therefore be less likely to be degraded
XX Sequence 190 AA;

Alignment Scores:
Pred. No.: 9.03e-15 Length: 190
Score: 204.50 Matches: 59
Percent Similarity: 47.89% Conservative: 43
Best Local Similarity: 27.70% Mismatches: 84
Query Match: 13.12% Indels: 27
DB: 3 Gaps: 5

US-09-647-661-1 (1-900) x AAY90339 (1-190)

Qy 151 ATGAATTTTGGATCGAGAAAAGAGAACACTATTAACGAGCGCATTTCTTGAAG 210
Db 1 MetThrGlnLeuThrArgGluGlnValLeuGluLeuPheHisGlnArgSerSerThrArg 20
```

```
Qy 211 ATGTTTGACAGCCATTATGAGTTTCTAGTGAAGAATTAGAAGAAATCGCTGAATCGCC 270
Db 21 TyrTyrAspProThrLysLysIleSerAspGluAspPheGluCysIleLeuGluCysGly 40
Qy 271 AGGCTTATCGCAAGCTCTTACAACAGCAGCCATGGCATTGTTGTCATGGTTACTAATAG 330
Db 41 ArgLeuSerProSerValGlySerGluProIrpLysPheLeuValIleGlnAsnLys 60
Qy 331 GATTTAATAAACAATTCGAGTCACAGCTACTTAAATGAAGAAATGATTAAAGCCGCT 390
Db 61 ThrLeuArgGluLysMetLysProPheSerTrpGlyMetIleAenGlnLeuAspAsnCys 80
Qy 391 TCAGCGTTAATGCTGCTATGCTCTTAAAGACTAGCAGATTGTTACACACGGCCATTAC 450
Db 81 SerHisLeuValValIleLeuAlaLysLys-----AsnAlaArgTyr 94
Qy 451 -----ATGCAAAACCTTTACCCGAGTCTTATAGTTAGGTAGAGTATCCCTTTTTCCT 504
Db 95 AspSerGlnGlnAlaAlaLeuThrLysTyrLysAla-----Leu 108
Qy 505 CAAATGCTTGGCTGAGATTCAACACAGCATGCAAGATTAGAAAGCTATATTTAGAG 564
Db 109 GlnGluGluAspMetLysLeuLeuGluAsnAspArgThrLeuPheAspTrpCysSerLys 128
Qy 565 CAATGCTATATCGCTGGGGCAAAATTTGCATGGCGTCAGCTTAATCGGATTGGATAGT 624
Db 129 GlnThrTyrIleAlaLeuAlaAsnMetLeuThrGlyAlaSerAlaLeuGlyLeuAspSer 148
Qy 625 TGCATTATGAGGCTTTGATCCCTTAAAGTGGTGAAAGTTTAA---GAAGAGCGTATC 681
Db 149 CysProIleGluGlyPheHisTyrAspLysMetAsnGluCysLeuAlaGluGluGlyLeu 168
Qy 682 AATAAGCCTTAAATCGCATGCTTGTATCGCTTTGGCAAGAGGTTGGCAGAGCGAGCCAA 741
Db 169 PheAspProGln-----GlutyrAlaVal 176
Qy 742 AAATCAAGAAATCAAAAGTTGATCGGATTACTTGGTTG 780
Db 177 LysSerArgLysGlyLeuAspGluValValLysTrpVal 189

RESULT 11
ID AAY90344 standard; protein; 201 AA.
XX AAY90344;
AC AAY90344;
DT 22-NOV-2000 (first entry)
DE C. jejuni nitroreductase.
XX Nitroreductase; CB1954; cytotoxic 4-hydroxylamine; 4HX derivative;
KW monofunctional alkylating agent; prodrug therapy; tumour therapy;
KW non-cytotoxic 2-hydroxylamine derivative; enzyme.
XX Campylobacter jejuni.
OS WO200047725-A1.
PN 17-AUG-2000.
PD 10-FEB-2000; 2000WO-GB0000431.
PF 10-FEB-1999; 99GB-00003019.
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX Minton N, Anlezark G, Vaughan T;
XX WPI; 2000-558214/51.
XX N-PSDB; AAA37746.
PT A nitroreductase useful for producing a medicament for anti-tumour therapy
PT preferentially reduces CB1954 to a cytotoxic 4-hydroxylamine derivative
```


QY 493 CCTCTCTTTGCTCAAAATGCTGGTGGAGATTCACACACAGCATGCAAGATTAGAA--- 549
 Db 114 ---LysAlaAlaAsnAspLysGlyArgLysPheSerAlaAspMetHisArgLysAspLeu 132
 QY 550 -----ACGTATATTTAGAGCAATGCTATATCGCTGCTGGGGCAAAATTTCATG 597
 Db 133 HisAspAspAlaGluTrpMetAlaLysGlnValTyrLeuAsnValGlyAsnPheLeuLeu 152
 QY 598 GGCGTGAAGCTTAAGGATTGGATGATGATTCATTATGGAGCTTTGAT 645
 Db 153 GlyValAlaAlaLeuGlyLeuAspAlaValProIleGluGlyPheAsp 168

RESULT 13
 AAW13083
 ID AAW13083 standard; protein; 217 AA.
 XX AC AAW13083;
 XX DT 08-MAY-1997 (first entry)
 XX XX Escherichia coli flavin reductase.
 XX XX
 XX XX Flavin reductase; nitro reductase; recombinant production; thermostable;
 XX KW high activity; random mutation.
 XX XX
 XX XX Escherichia coli.
 XX XX Location/Qualifiers
 FH Key
 FT Misc-difference 124 /note= "wild type Phe substituted with Ser"
 FT
 XX JF090000266-A.
 XX 07-JAN-1997.
 XX 19-JUN-1995; 95JP-00176736.
 XX PR 19-JUN-1995; 95JP-00176736.
 XX XX (CHCC) CHISSO CORP.
 XX WPI; 1997-112849/11.
 XX DR N-PSDB; AAT61841.
 XX XX
 XX XX Escherichia coli flavin reductase - prepd. by mutating wild type E. coli
 PT nitro reductase gene.
 XX Claim 7; Page 7; 9pp; Japanese.
 XX
 XX The present sequence is an Escherichia coli flavin reductase (FR), which
 CC was prepared by mutating an E. coli derived wild type nitro reductase
 CC (NR) gene. The FR gene can be used for the recombinant production of the
 CC new stable, high activity FR. The recombinant plasmid pNR1DNA was used as
 CC the template for the random mutation of the E. coli NR gene by PCR. E.
 CC coli containing a plasmid capable of expressing the mutated NR gene were
 CC prepared, and an E. coli containing a plasmid expressing the FR gene
 CC selected. The FR encoded by pNR247 was purified, found to be stable at
 CC temperatures of up to 40 degrees C and the base sequence of its gene
 CC determined
 XX
 XX SQ Sequence 217 AA;
 Alignment Scores:
 Pred. No.: 2,45e-13 Length: 217
 Score: 192.50 Matches: 49
 Percent Similarity: 48.30% Conservative: 36
 Best Local Similarity: 27.84% Mismatches: 50
 Query Match: 12.35% Indels: 41
 DB: 2 Gaps: 5
 US-09-647-661-1 (1-900) x AAW13083 (1-217)
 QY 193 GAGCGCCATCTTCGCAAGATGTTGACGCCATTATGAGTTTCTAGTGAGAAATTAGAA 252

Db 9 LysArgHisSerThrLysAlaPheAspAlaSerLysLysLeuThrProGluGlnAlaGlu 28
 QY 253 GAATCGCTGAAATCGCCAGGCTATCGCAAGCTCTTACACACGCGCCATTCGCTTT 312
 Db 29 GlnIleLysThrLeuGlnTyrSerProSerSerThrAsnSerGlnProItrPhePhe 48
 QY 313 GTGATGGTTACTAATAAGGATTTAAAAAACAATTCAGTGCACAGC----- 360
 Db 49 IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68
 QY 361 TACTTTAATGAAGAAATGATTAAAGCGCTTCACGCTTAATGGTGGTATGCTCTTTAAGA 420
 Db 69 ValPheAsnGluArgLysMetLeuAspAlaSerHisValValValPheCysAlaLysThr 88
 QY 421 CCTAGCGAGTTGTTACCACACAGCGCATTTACATGCAAAACCTTTAC----- 465
 Db 89 Ala-----MetAspAspValTTrpLeuLysLeuValVal 99
 QY 466 -----CCGAGTCTTTATAAGTTAGGTAGATGATC 492
 Db 100 AspGlnGluAspAlaAspGlyArgPheAlaThrProGluAla----- 113
 QY 493 CCTCTCTTTGCTCAAAATGCTGGCGTGAGATTCACACACAGCATGCAAGATTAGAA--- 549
 Db 114 ---LysAlaAlaAsnAspLysGlyArgLysPheSerAlaAspMetHisArgLysAspLeu 132
 QY 550 -----ACGTATATTTAGAGCAATGCTATATCGCTGCTGGGGCAAAATTTCATG 597
 Db 133 HisAspAspAlaGluTrpMetAlaLysGlnValTyrLeuAsnValGlyAsnPheLeuLeu 152
 QY 598 GGCGTGAAGCTTAATGGGATTGGATGATTCATTATGGAGCTTTGAT 645
 Db 153 GlyValAlaAlaLeuGlyLeuAspAlaValProIleGluGlyPheAsp 168

RESULT 14
 ABP70750
 ID ABP70750 standard; protein; 217 AA.
 XX AC ABP70750;
 XX DT 15-MAY-2003 (first entry)
 XX XX Mutant E. coli nitroreductase, NFSB, Y686/F124Q.
 DE Nitroreductase; enzyme; cytostatic; NFSB; dihydropteridine reductase;
 KW oxygen-insensitive NAD(P)H nitroreductase; NTR; NFSB; prodrug CD1954;
 KW anti-tumour; cancer; mutant; mutain.
 XX
 XX Escherichia coli.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 68
 FT Misc-difference 124 /note= "Wild-type Tyr substituted by Gly"
 FT Misc-difference 124 /note= "Wild-type Phe substituted by Gln"
 XX
 XX WO2003018788-A2.
 XX PD 06-MAR-2003.
 XX XX 21-AUG-2002; 2002WO-GB003833.
 XX PR 21-AUG-2001; 2001GB-00020394.
 XX PR 06-SEP-2001; 2001GB-00021662.
 XX PR 03-OCT-2001; 2001US-0326846P.
 XX PA (MLML-) ML LAB PLC.
 XX XX Grove JI, Searle PF, Lovering AL;
 XX WPI; 2003-300732/29.
 DR

XX New recombinant mutant nitroreductase having increased nitroreductase
PT activity, useful as a medicament for treating cancer, for converting a
PT prodrug into a cytotoxic agent, and in designing or screening for
PT improved prodrugs.
XX
XX
XX Claim 15; Page; 63pp; English.
XX
XX The present invention relates to mutant nitroreductases (ABP70750-
CC ABP70759), which have increased nitroreductase activity as compared to
CC the wild-type enzyme, E. coli NFBS nitroreductase (EC 1.6.99.7, also
CC known as oxygen-insensitive NAD(P)H nitroreductase or dihydropteridine
CC reductase (NTR)). NFBS reduces the aziridinyl prodrug CD1954 within cells
CC to produce a genotoxic 4-hydroxylamine derivative and so is useful as an
CC anti-tumour agent. The mutant nitroreductases are therefore useful as
CC medicaments for treating cancer. Note: The present sequence is not shown
CC in the specification but is derived from the Escherichia coli wild-type
CC nitroreductase sequence given in figure 9 (see ABP70750)

XX SQ Sequence 217 AA;

Alignment Scores:
Pred. No.: 2,45e-13 Length: 217
Score: 192.50 Matches: 50
Percent Similarity: 47.16% Conservatives: 33
Best Local Similarity: 28.41% Mismatches: 52
Query Match: 12.35% Indels: 41
DB: 6 Gaps: 5

US-09-647-661-1 (1-900) X ABP70750 (1-217)

QY 193 GAGGCCATTCTTCAGAGATGTTTGACGCCATTATGAGTTTCTAGTGAAGAAATTAGAA 252
Dy 9 LysArgHisSerThrLysAlaPheAspAlaSerLysLysLeuThrProGluGlnAlaGlu 28
QY 253 GAATTCGCTGAATCGCCAGCTATCGCAAGCTTTACACAGCGAGCCATGCGATTTT 312
Dy 29 GlnIleLysThrLeuLeuGlnTyrSerProSerSerThrAsnSerGlnProTrpHisPhe 48
QY 313 GTGATGGTTACTAATAGGATTTAA--AAACAATGTCAGTGCACAGC 360
Dy 49 IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyAsnGly 68
QY 361 TACTTTAATGAAGAATGATTAAGCGCTTCAGCGTTAATGGTGTATGCTCTTTAAGA 420
Dy 69 ValPheAsnGluArgLysMetLeuAspAlaSerHisValValPheCysAlaLysThr 88
QY 421 CCTAGCGAGTTGTTACCACAGCGCCATTACATGCAAAACCTTTAC----- 465
Dy 89 Ala-----MetAspAspValTrpLeuLysLeuValVal 99
QY 466 -----CCGAGTCTTATAGGTTAGTGATC 492
Dy 100 AspGlnGluAspAlaAspGlyArgPheAlaThrProGluAla----- 113
QY 493 CTTCTTTTGTCAATGCTTGGCGTGAGATTCAACACAGCATGCAAGATTAGAA--- 549
Dy 114 ---LysAlaAlaAsnAspLysGlyArgLysPheGlnAlaAspMetHisArgLysAspLeu 132
QY 550 -----ACGTATATTTAGACGAATGCTATATCCGTGGGGCAAAATTGCGATG 597
Dy 133 HisAspAspAlaGluTrpMetAlaLysGlnValTrpLeuAsnValGlyAsnPheLeuLeu 152
QY 598 GCGGTGAGCTTAATGGGATGAGTAGTTGCAATTATGGAGCTTTGAT 645
Dy 153 GlyValAlaAlaLeuGlyLeuAspAlaValProIleGluGlyPheAsp 168

RESULT 15

ABP70759

ID ABP70759 standard; protein; 217 AA.

XX AC

XX ABP70759;

XX

DT 15-MAY-2003 (first entry)
XX
DE Mutant E.coli nitroreductase, NFBS, #7.
XX
KW Nitroreductase; enzyme; cytostatic; NFBS; dihydropteridine reductase;
KW oxygen-insensitive NAD(P)H nitroreductase; NTR; NFBS; prodrug CD1954;
KW anti-tumour; cancer; mutant; mutuin.
XX
OS Escherichia coli.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 124
FT /note= "Wild-type Phe can be substituted by Asn, Cys,
FT Gly, Lys, Met, Trp, Tyr, Ala, Gln, His, Ile, Leu, Ser,
FT Thr or Val"
XX
PN WO2003018788-A2.
XX
PD 06-MAR-2003.
XX
XX 21-AUG-2002; 2002WO-GB0038333.
XX
PR 21-AUG-2001; 2001GB-00020294.
PR 06-SEP-2001; 2001GB-00021662.
PR 03-OCT-2001; 2001US-0326846P.
XX
XX (MLML-) ML LAB PLC.
XX
XX Grove JI, Searle PF, Lovering AJ;
XX
XX WPI; 2003-300732/29.
XX
PT New recombinant mutant nitroreductase having increased nitroreductase
PT activity, useful as a medicament for treating cancer, for converting a
PT prodrug into a cytotoxic agent, and in designing or screening for
PT improved prodrugs.
XX
XX Claim 23; Page; 63pp; English.
XX
PS The present invention relates to mutant nitroreductases (ABP70750-
CC ABP70759), which have increased nitroreductase activity as compared to
CC the wild-type enzyme, E. coli NFBS nitroreductase (EC 1.6.99.7, also
CC known as oxygen-insensitive NAD(P)H nitroreductase or dihydropteridine
CC reductase (NTR)). NFBS reduces the aziridinyl prodrug CD1954 within cells
CC to produce a genotoxic 4-hydroxylamine derivative and so is useful as an
CC anti-tumour agent. The mutant nitroreductases are therefore useful as
CC medicaments for treating cancer. Note: The present sequence is not shown
CC in the specification but is derived from the Escherichia coli wild-type
CC nitroreductase sequence given in figure 9 (see ABP70750)

XX SQ Sequence 217 AA;

Alignment Scores:
Pred. No.: 4.21e-13 Length: 217
Score: 190.50 Matches: 49
Percent Similarity: 47.73% Conservatives: 35
Best Local Similarity: 27.84% Mismatches: 51
Query Match: 12.22% Indels: 41
DB: 6 Gaps: 5

US-09-647-661-1 (1-900) X ABP70759 (1-217)

QY 193 GAGGCCATTCTTCAGAGATGTTTGACGCCATTATGAGTTTCTAGTGAAGAAATTAGAA 252
Dy 9 LysArgHisSerThrLysAlaPheAspAlaSerLysLysLeuThrProGluGlnAlaGlu 28
QY 253 GAATTCGCTGAATCGCCAGCTATCGCAAGCTTTACACAGCGATGCGATTTT 312
Dy 29 GlnIleLysThrLeuLeuGlnTyrSerProSerSerThrAsnSerGlnProTrpHisPhe 48
QY 313 GTGATGGTTACTAATAGGATTTAA-----CAAAATGTCAGTGCACAGC----- 360

Db 49 IleValAlaSerThrGluGluGlyLysAlaArgValAlaLysSerAlaAlaGlyAsnTyr 68
 Qy 361 TACTTTTAATCAAGAAATGATTAAAAAGCCCTTCAGCGTTAATGTGGTATGCTCTTTAAGA 420
 Db 69 ValPheAsnGluArgLysMetLeuAspAlaSerHisValValValPheCysAlaLysThr 88
 Qy 421 CCTAGCGAGTGTTCACACACGCGCATTCACATGCGAAACCTTTAC----- 465
 Db 89 Ala-----MetAspAspValTrpLeuLysLeuValVal 99
 Qy 466 -----CCGGAGTCTTATAAGTTAGGTTAGATGATC 492
 Db 100 AspGlnGluAspAlaAspGlyArgPheAlaThrProGluAla----- 113
 Qy 493 CCTCTCTTTGCTCAAAATGCTTGGCGTGAAGATTCAACACACAGCATGCAAAAGATTAGAA----- 549
 Db 114 ---LysAlaAlaAsnAspLysGlyArgLysPhe***AlaAspMetHisArgLysAspLeu 132
 Qy 550 -----AGCTATATTTTATGAGCAATGCCTATATCGCTGTGGGGCAAAATTTGCATG 597
 Db 133 HisAspAspAlaGluTrpMetAlaLysGlnValTyrLeuAsnValGlyAsnPheLeuLeu 152
 Qy 598 GCGGTGAGCTTAATGGGATTGATAGTTGCATTATTGGAGGCTTTGAT 645
 Db 153 GlyValAlaLeuGlyLeuAspAlaValProIleGluGlyPheAsp 168

Search completed: April 20, 2004, 19:24:49
Job time : 64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 20, 2004, 19:11:02 ; Search time 64 Seconds
(without alignments)
7753.307 Million cell updates/sec

Title: US-09-647-661-1

Perfect score: 1559

Sequence: 1 tgcagaattttcacagagc.....gggggtgtttcaagcgtttc 900

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Fgapop 10.0 , Fgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 2249750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame-n2p.model -DB=us-09-647-661-1
-O=/cg2_1/USPTO_spool_p/US09647661/runat_20042004_135346_2594/app_query.fasta_1.1095
-DB=Published Applications AA -GEMT=fastan -SUFFIX=xapb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct-THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MOD=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09647661 @CGN 1 1 13 @runat_20042004_135346_2594
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

- 1: /cg2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cg2_6/ptodata/1/pubpaa/FCI_NEW_PUB.pep.*
- 3: /cg2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cg2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cg2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cg2_6/ptodata/1/pubpaa/ACTUS_PUBCOMB.pep.*
- 7: /cg2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cg2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cg2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cg2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cg2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cg2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cg2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cg2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cg2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cg2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cg2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cg2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

RESULT 1

US-09-882-227-180

Sequence 180, Application US/09882227

Publication No. US20030158396A1

GENERAL INFORMATION:

APPLICANT: Kleantous, Harold

APPLICANT: Al-Garawi, Amal

APPLICANT: Miller, Charles

APPLICANT: Tomb, Jean-Francois

APPLICANT: Coomen, Raymond P.

TITLE OF INVENTION: Identification of Polynucleotides

TITLE OF INVENTION: Encoding No. US20030158396A1

TITLE OF INVENTION: Genome

FILE REFERENCE: 06132/047002

CURRENT APPLICATION NUMBER: US/09/882,227

CURRENT FILING DATE: 2001-06-15

PRIOR APPLICATION NUMBER: US 08/902,615

PRIOR FILING DATE: 1997-07-29

NUMBER OF SEQ ID NOS: 638

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 180

LENGTH: 210

TYPE: PRT

ORGANISM: Helicobacter pylori

ALIGNMENTS

1	1072	68.8	210	10	US-09-882-227-180	Sequence 180, App
2	1067	68.4	210	15	US-10-364-397-29	Sequence 29, Appl
3	1041	66.8	210	12	US-10-335-977-6487	Sequence 6487, Ap
4	636	40.8	134	12	US-10-335-977-6486	Sequence 6486, Ap
5	626	40.2	127	12	US-10-335-977-6485	Sequence 6485, Ap
6	448	28.7	97	12	US-10-335-977-6484	Sequence 6484, Ap
7	427	27.4	91	12	US-10-335-977-6483	Sequence 6483, Ap
8	246	15.8	284	9	US-09-895-913A-102	Sequence 102, App
9	229	14.7	283	12	US-10-335-977-5526	Sequence 5526, Ap
10	218.5	14.0	220	15	US-10-364-397-17	Sequence 17, Appl
11	198.5	12.7	201	15	US-10-364-397-23	Sequence 23, Appl
12	190	12.2	217	12	US-10-335-977-6413	Sequence 6413, Ap
13	185.5	11.9	202	15	US-10-364-397-12	Sequence 12, Appl
14	182	11.2	221	14	US-10-169-048-2	Sequence 2, Appl
15	174	11.1	200	15	US-10-364-397-19	Sequence 19, Appl
16	172.5	11.1	209	15	US-10-364-397-21	Sequence 21, Appl
17	153	9.8	174	15	US-10-364-397-21	Sequence 21, Appl
18	150	9.6	198	10	US-09-769-744A-50	Sequence 30, Appl
19	136.5	8.8	205	15	US-10-364-397-18	Sequence 18, Appl
20	124	8.0	172	15	US-10-364-397-20	Sequence 20, Appl
21	116	7.4	185	12	US-10-335-977-6272	Sequence 6272, Ap
22	114	7.3	197	12	US-10-282-122A-53487	Sequence 53487, A
23	113	7.2	205	14	US-10-158-761-8103	Sequence 8103, Ap
24	108.5	6.8	79	12	US-10-335-977-6411	Sequence 6411, Ap
25	105.5	6.8	219	12	US-10-282-122A-61643	Sequence 61643, A
26	102.5	6.6	289	12	US-10-236-417-90	Sequence 90, Appl
27	97.5	6.3	90	12	US-10-335-977-6412	Sequence 6412, Ap
28	93.5	6.0	235	14	US-10-153-382-7	Sequence 7, Appl
29	92.5	5.9	218	12	US-10-282-122A-53372	Sequence 53372, A
30	92	5.9	234	9	US-09-800-729-150	Sequence 150, App
31	92	5.9	234	11	US-09-833-245-2210	Sequence 2210, Ap
32	91.5	5.9	241	14	US-10-221-945-1	Sequence 1, Appl
33	91	5.8	1125	12	US-10-282-122A-47213	Sequence 47213, A
34	90.5	5.8	214	14	US-10-229-567-11	Sequence 11, Appl
35	89.5	5.7	215	15	US-10-307-724-122	Sequence 122, App
36	89	5.7	1001	12	US-10-335-977-5785	Sequence 5785, Ap
37	89	5.7	1001	12	US-10-335-977-5786	Sequence 5786, Ap
38	88.5	5.7	141	14	US-10-153-382-23	Sequence 23, Appl
39	88.5	5.7	214	12	US-10-411-037-55	Sequence 55, Appl
40	88.5	5.7	214	12	US-10-411-026-55	Sequence 55, Appl
41	88.5	5.7	237	9	US-09-056-160B-100	Sequence 100, App
42	88.5	5.7	237	14	US-10-234-671-100	Sequence 100, App
43	88.5	5.7	491	13	US-10-011-125-2	Sequence 2, Appl
44	88	5.6	625	12	US-10-282-122A-46895	Sequence 46895, A
45	87.5	5.6	214	12	US-10-364-953-1	Sequence 1, Appl

08/19/2011/1947

US-09-882-227-180

Alignment Scores:

Pred. No.: 1.59e-110 Length: 210
 Score: 1072.00 Matches: 207
 Percent Similarity: 99.05% Conservativeness: 1
 Best Local Similarity: 98.57% Mismatches: 2
 Query Match: 68.78% Indels: 0
 DB: 10 Gaps: 0

US-09-647-661-1 (1-900) x US-09-882-227-180 (1-210)

QY 151 ATCAAAATTTTGGATCAGGAAAGAAAGACAACTATTAAACGAGCGCCATTCTTGCAG 210
 Db 1 MetLysPheLeuAspGlnGlyArgGlnLeuLeuAsnGluArgHisSerCysLys 20
 QY 211 ATGTTTGACGCCATTATGAGTTTCTAGTGAAGAAATTAGAAGAAATCGCTCAATGCC 270
 Db 21 MetPheAspSerHisTyrGluPheSerThrGluLeuGluGluAlaGluAla 40
 QY 271 AGGCTATGCCCAAGCTCTTACACACGCGCATGGCATTGTGATGGTTACTAATAAG 330
 Db 41 ArgLeuSerProSerSerTyrAsnThrGlnProThrHisPheValMetValThrAspLys 60
 QY 331 GATTTTAAAAACAAATTCAGTGCACAGCTACTTTAATGAAGAAATGATTAAGAGCGCT 390
 Db 61 AspLeuLysLysGlnIleAlaAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
 QY 391 TCAGCGTTAATGGTGTGCTCTTTAAGACTAGCGAGTGTGTACACACGCGCATTAC 450
 Db 81 SerAlaLeuMetValCysSerLeuArgProSerGluLeuLeuProHisGlyHisTyr 100
 QY 451 ATGCAAAACCTTTACCCGGAGTCTTATAAGTTAGAGTGTACCTCTTTTGTCTCAATG 510
 Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
 QY 511 CTGCGGTGAGATTCAACACAGCATGCAAGATTAGAAGCTATATTATTAGACCAATGC 570
 Db 121 LeuGlyValArgPheAsnHisSerMetGlnArgLeuGluSerTyrIleLeuGluGlnCys 140
 QY 571 TATATCGCTGTGGGCAAAATTTGCATGGCGGTGAGCTTAATGGGATTGGATTGCTATT 630
 Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
 QY 631 ATGAGGCTTTGATCCTTTAAAGTGGTGAAGTTTAGAGAGCGTATCAATAAGCTT 690
 Db 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluGluArgIleAsnLysPro 180
 QY 691 AAAATCGCATGCTTGCATCGCTTTGGCAAGAGGGTGGCAGAGCGCAAAATCAAGA 750
 Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
 QY 751 AAATCAAAAGTTGATCGGATTACTTGGTTG 780
 Db 201 LysSerLysValAspAlaIleThrTrpLeu 210

RESULT 2

US-10-364-397-29

; Sequence 29, Application US/10364397
 ; Publication No. US20040014191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Minton, Nigel
 ; APPLICANT: Anlezark, Gill
 ; APPLICANT: Vaughan, Thomas
 ; TITLE OF INVENTION: Nitroreductase Enzymes
 ; FILE REFERENCE: 1581.0850001
 ; CURRENT APPLICATION NUMBER: US/10/364,397
 ; CURRENT FILING DATE: 2003-02-12
 ; PRIOR APPLICATION NUMBER: 09/913,068
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00431
 ; PRIOR FILING DATE: 2000-02-10
 ; PRIOR APPLICATION NUMBER: UK 9903019.9

; PRIOR FILING DATE: 1999-02-10

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 29

; LENGTH: 210

; TYPE: PRT

; ORGANISM: Helicobacter pylori

US-10-364-397-29

Alignment Scores:

Pred. No.: 5.73e-110 Length: 210
 Score: 1067.00 Matches: 206
 Percent Similarity: 99.05% Conservativeness: 2
 Best Local Similarity: 98.10% Mismatches: 2
 Query Match: 68.44% Indels: 0
 DB: 15 Gaps: 0

US-09-647-661-1 (1-900) x US-10-364-397-29 (1-210)

QY 151 ATCAAAATTTTGGATCAGGAAAGAAAGACAACTATTAAACGAGCGCCATTCTTGCAG 210
 Db 1 MetLysPheLeuAspGlnGlyArgGlnLeuLeuAsnGluArgHisSerCysLys 20
 QY 211 ATGTTTGACAGCCATTATGAGTTTCTAGTGAAGAAATTAGAAGAAATCGCTCAATGCC 270
 Db 21 MetPheAspSerHisTyrGluPheSerThrGluLeuGluGluAlaGluAla 40
 QY 271 AGGCTATGCCCAAGCTCTTACACACGCGCATGGCATTGTGATGGTTACTAATAAG 330
 Db 41 ArgLeuSerProSerSerTyrAsnThrGlnProThrHisPheValMetValThrAspLys 60
 QY 331 GATTTTAAAAACAAATTCAGTGCACAGCTACTTTAATGAAGAAATGATTAAGAGCGCT 390
 Db 61 AspLeuLysLysGlnIleAlaAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
 QY 391 TCAGCGTTAATGGTGTGCTCTTTAAGACTAGCGAGTGTGTACACACGCGCATTAC 450
 Db 81 SerAlaLeuMetValCysSerLeuLysProSerGluLeuLeuProHisGlyHisTyr 100
 QY 451 ATCAAAACCTTTACCCGGAGTCTTATAAGTTAGAGTGTACCTCTTTTGTCTCAATG 510
 Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120
 QY 511 CTGCGGTGAGATTCAACACAGCATGCAAGATTAGAAGCTATATTATTAGACCAATGC 570
 Db 121 LeuGlyValArgPheAsnHisSerMetGlnLysLeuGluSerTyrIleLeuGluGlnCys 140
 QY 571 TATATCGCTGTGGGCAAAATTTGCATGGCGGTGAGCTTAATGGGATTGGATTGCTATT 630
 Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
 QY 631 ATGAGGCTTTGATCCTTTAAAGTGGTGAAGTTTAGAGAGCGTATCAATAAGCTT 690
 Db 161 IleGlyGlyPheAspProLeuLysValGlyGluValLeuGluGluArgIleAsnLysPro 180
 QY 691 AAAATCGCATGCTTGCATCGCTTTGGCAAGAGGGTGGCAGAGCGCAAAATCAAGA 750
 Db 181 LysIleAlaCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerGlnLysSerArg 200
 QY 751 AAATCAAAAGTTGATCGGATTACTTGGTTG 780
 Db 201 LysSerLysValAspAlaIleThrTrpLeu 210

RESULT 3

US-10-335-977-6487

; Sequence 6487, Application US/10335977
 ; Publication No. US20040052799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DOUGLAS SMITH et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; RELATING TO HELICOBACTER PYLORI FOR
 ; DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 10031

; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6487:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...210
; SEQUENCE DESCRIPTION: SEQ ID NO: 6487:
US-10-335-977-6487

Alignment Scores:
Pred. No.: 4.58e-107 Length: 210
Score: 1041.00 Matches: 200
Percent Similarity: 97.14% Conservative: 4
Best Local Similarity: 95.24%
Query Match: 66.77% Indels: 0
DB: 12 Gaps: 0

US-09-647-661-1 (1-900) x US-10-335-977-6487 (1-210)

QY 151 ATGAATTTTGGATCAGGAAAAAGAGACAACTATTAAACGAGCGCCATTCTTGCAG 210
Db 1 MetLysPheLeuAspHisGluLysArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTTGACGCAATATAGTTTCTAGTGAAGATTAGAAGAAATCGCTGAATCGCC 270
Db 21 MetPheAspSerHisTyrGluPheSerSerThrGluLeuGluGluLeuAlaGluLeuAla 40
QY 271 AGGCTATCGCAAGCTTTACACACGACGACGATGTCATTTGTGATGTTACTAATAAG 330
Db 41 ArgLeuSerProSerSerTyrAsnThrGlnProThrPhePheValMetValThrAsnLys 60
QY 331 GATTAAAAAACAATTCAGTCGACGCTACTTTAATGAAGAAATGATTAAAGCGCT 390
Db 61 AspLeuLysGlnIleAlaAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 391 TCAGCGTTAATGGTGTATGCTCTTTAAGACCTAGCGAGTGTGTACCAACGCGCATAC 450
Db 81 SerAlaLeuMetValValCysSerLeuLysProSerGluLeuLeuProThrSerHisTyr 100
QY 451 ATCAAAACCTTTACCGGAGCTTATAGGTTAGAGTGCATCCCTTCTTTCTCAATG 510
Db 101 MetGlnAsnLeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMet 120

QY 511 CTTGGCGTGAGATTCAACACAGCATGCAAGATTAGAAAGCTATATTTAGAGCAATGC 570
Db 121 LeuGlyValArgPheAsnHisSerMetGlnLysLeuGluSerTyrIleLeuGluGlnCys 140
QY 571 TATATCGCTGTGGGCAAAATTTGCATGGCGGTGAGCTTAATGGGATTGGATTTGCATT 630
Db 141 TyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIle 160
QY 631 ATTGGAGCGCTTTGATCCTTTAAAGTGGGTGAAGTTTATAGAAGAGCGGTATCAATAAGCCT 690
Db 161 IleGlyGlyPheAspProLeuLysValGlyGluIleLeuGluGluArgIleAsnLysPro 180
QY 691 AAAATCGCATGCTTGTATCGCTTTGGCAAGAGAGGTGGCAGACGAGCCAAATAATCAAGA 750
Db 181 LysIleValCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerLysLysSerArg 200
QY 751 AAATCAAAAGTTGATCGATTACTTGTGTTG 780
Db 201 LysSerLysValAspAlaIleThrProLeu 210

RESULT 4

US-10-335-977-6486
; Sequence 6486, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6486:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...134
; SEQUENCE DESCRIPTION: SEQ ID NO: 6486:
US-10-335-977-6486

Alignment Scores:
Pred. No.: 6.47e-62 Length: 134

Score: 636.00 Matches: 122
Percent Similarity: 95.52% Conservative: 6
Best Local Similarity: 91.04% Mismatches: 6
Query Match: 40.80% Indels: 0
DB: 12 Gaps: 0

US-09-647-661-1 (1-900) x US-10-335-977-6486 (1-134)

QY 379 ATTAAAGCGCTTACGGTAAATGGTGTATGCTTTTAAGACCTAGCGAGTTGTACCA 438
: : : : :
Db 1 LeuYsAlaLeuGlnLeuLeuMetValValCysSerLeuYsProSerGluLeuLeuPro 20
QY 439 CACGGCATTACATGCAAAACCTTTACCGGAGCTTATAGGTTAGAGTATCCCTTCT 498
: : : : :
Db 21 ThrSerHisTyrMetGlnAsnLeuYsProGluSerTyrLysValArgValIleProSer 40
QY 499 TTTGCTCAATGCTTGGCGTGAAGTCAACACAGCATGCAAAAGATTAGAAAGCTATAT 558
: : : : :
Db 41 PheAlaGlnMetLeuGlyValArgPheAsnHisSerMetGlnLysLeuGluSerTyrIle 60
QY 559 TTAGACATGCTATATCGCTGTGGGCAAAATTTGATGGCGGTGAGCTTATGGATG 618
: : : : :
Db 61 LeuGluGlnCysTyrIleAlaValGlyGlnIleCysMetGlyValSerLeuMetGlyLeu 80
QY 619 GATAGTTGCTATTATGGAGCTTTGATCCTTTAAAGTGGGTGAAGTTTGAAGAGCGT 678
: : : : :
Db 81 AspSerCysIleIleGlyGlyPheAspProLeuLysValGlyGluIleLeuGluArg 100
QY 679 ATCAATAGCCTAAATCGCATGCTTGTGATCGCTTTGGCAAGAGGTGCGAAGCGAGC 738
: : : : :
Db 101 IleAsnLysProLysIleValCysLeuIleAlaLeuGlyLysArgValAlaGluAlaSer 120
QY 739 CAAAATCAAGAAATCAAAAGTTGATCGATTACTTGTG 780
: : : : :
Db 121 LysLysSerArgLysSerLysValAspAlaIleThrIleLeu 134

RESULT 5
US-10-335-977-6485
; Sequence 6485, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6485:
; SEQUENCE CHARACTERISTICS:

LENGTH: 127 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...127
SEQUENCE DESCRIPTION: SEQ ID NO: 6485:
US-10-335-977-6485

Alignment Scores:
Pred. No.: 8,33e-61 length: 127
Score: 626.00 Matches: 120
Percent Similarity: 97.64% Conservative: 4
Best Local Similarity: 94.49% Mismatches: 3
Query Match: 40.15% Indels: 0
DB: 12 Gaps: 0

US-09-647-661-1 (1-900) x US-10-335-977-6485 (1-127)

QY 400 ATGTGCTATGCTCTTTTAAGACCTAGCGAGTTGTACCAACGCGCCATTACATGCAAAAC 459
: : : : :
Db 1 MetValValCysSerLeuYsProSerGluLeuLeuProThrSerHisTyrMetGlnAsn 20
QY 460 CTTTACCGGAGTCTTAAAGTTAGAGTATCCCTTCTTTTGTCTCAAATGCTTGGCGTG 519
: : : : :
Db 21 LeuTyrProGluSerTyrLysValArgValIleProSerPheAlaGlnMetLeuGlyVal 40
QY 520 AGATTCAACACAGCATGCAAAAGATTAGAAAGCTATATTTTAGAGCAATGCTATATCGCT 579
: : : : :
Db 41 ArgPheAsnHisSerMetGlnLysLeuGluSerTyrIleLeuGluGlnCysTyrIleAla 60
QY 580 GTGGGGCAAAATTTGCTATGCGGCGTGAAGTTTAAAGAGCGGTATCAATAGCCTTAAATCGCA 639
: : : : :
Db 61 ValGlyGlnIleCysMetGlyValSerLeuMetGlyLeuAspSerCysIleIleGlyGly 80
QY 640 TTTGATCCTTTAAAGTGGGTGAAGTTTAAAGAGCGGTATCAATAGCCTTAAATCGCA 699
: : : : :
Db 81 PheAspProLeuLysValGlyGluIleLeuGluArgIleAsnLysProLysIleVal 100
QY 700 TGCTTGATCGCTTTGGCAAGAGGTGCGAAGCGAGCGCAAAATCAAGAAATCAAAA 759
: : : : :
Db 101 CysLeuIleAlaLeuGlyLysArgValAlaGluAlaSerLysLysSerArgLysSerLys 120
QY 760 GTTGCATCGATTACTTGGTTG 780
: : : : :
Db 121 ValAspAlaIleThrIleLeu 127

RESULT 6
US-10-335-977-6484
; Sequence 6484, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX

COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 6483:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...91
SEQUENCE DESCRIPTION: SEQ ID NO: 6483:
US-10-335-977-6483
Alignment Scores:
Pred. No.: 1,23e-38 Length: 91
Score: 427.00 Matches: 86
Percent Similarity: 95.60% Conservative: 1
Best Local Similarity: 94.51% Mismatches: 4
Query Match: 27.39% Indels: 1
DB: 12 Gaps: 0
US-09-647-661-1 (1-900) x US-10-335-977-6483 (1-91)
QY 151 ATGAATTTTGGATCAGGAAAGAAAGAACAACTATTAAACGAGCGCCATTCTTCCAAG 210
Db 1 MetLysPheLeuAspHisGluGluArgGlnLeuLeuAsnGluArgHisSerCysLys 20
QY 211 ATGTTTGACAGCCATTATGAGTTTCTAGTGAAGATTAGAGAAATCGCTGAATCGCC 270
Db 21 MetPheAspSerHisTyrGluPheSerSerThrGluLeuGluGluLeuAlaGluIleAla 40
QY 271 AGGCTATCGCAAGCTCTTACAAACAGCGCCATGCGCATTTTGTGATGGTTACTAATAAG 330
Db 41 ArgLeuSerProSerSerTyrAsnThrGlnProThrHisPheValMetValThrAsnLys 60
QY 331 GATTTAAAAACAAATTGCGAGTCGACAGCTACTTTAATGAGAAATGATTAAGCGCT 390
Db 61 AspLeuLysLysGlnIleAlaAlaHisSerTyrPheAsnGluGluMetIleLysSerAla 80
QY 391 TCAGC-GTTAATGCTGCTATGCTCTTTAAGACC 422
Db 81 SerAlaValAsnGlyMetLeuPheLysThr 91
RESULT 8
US-09-895-913A-102
Sequence 102, Application US/09895913A
Patent No. US20020160456A1
GENERAL INFORMATION:
APPLICANT: Kleanthous, Harold
APPLICANT: Al-Garawi, Amal
APPLICANT: Miller, Charles

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 6484:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...97
SEQUENCE DESCRIPTION: SEQ ID NO: 6484:
US-10-335-977-6484
Alignment Scores:
Pred. No.: 5.69e-41 Length: 97
Score: 448.00 Matches: 90
Percent Similarity: 94.85% Conservative: 2
Best Local Similarity: 92.78% Mismatches: 5
Query Match: 28.74% Indels: 1
DB: 12 Gaps: 0
US-09-647-661-1 (1-900) x US-10-335-977-6484 (1-97)
QY 133 AAAAAATAAGGAAATCAATGAAATTTTGGATCAGGAAAGAAAGAACAACTATTAAAC 192
Db 1 LysAsnLysGlyGluIleMetLysPheLeuAspHisGluGluArgGlnLeuLeuAsn 20
QY 193 GAGCGCCATCTTCAAGATGTTTCACAGCCATTATGAGTTTCTAGTGAAGATTAGAA 252
Db 21 GluArgHisSerCysLysMetPheAspSerHisTyrGluPheSerThrGluLeuGlu 40
QY 253 GAAATCGCTGAATCGCAGGCTATCGCAAGCTCTTACAAACAGCGCCATGCGCATTTT 312
Db 41 GluIleAlaGluIleAlaArgLeuSerProSerSerTyrAsnThrGlnProThrHisPhe 60
QY 313 GTGATGGTTACTAATAAGATTAAAAACAAATTCAGTGCACAGCTACTTTAATGAA 372
Db 61 ValMetValThrAsnLysAspLeuLysLysGlnIleAlaAlaHisSerTyrPheAsnGlu 80
QY 373 GAAATGATTAAAGCGCTTCAGC-GTTAATGCTGCTATGCTCTTTAAGACC 422
Db 81 GluMetIleLysSerAlaSerAlaValAsnGlyGlyMetLeuPheLysThr 97
RESULT 7
US-10-335-977-6483
Sequence 6483, Application US/10335977
Publication No. US20040052799A1
GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
City: Boston
STATE: Massachusetts

```
/ APPLICANT: Tomb, Jean Francois
/ APPLICANT: Comen, Raymond P.
/ TITLE OF INVENTION: Identification of Polynucleotides
/ TITLE OF INVENTION: Encoding No. US20020160454a1 Helicobacter Polypeptides in the
/ TITLE OF INVENTION: Genome
/ FILE REFERENCE: 06132/043002
/ CURRENT APPLICATION NUMBER: US/09/895,913A
/ PRIOR FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: US 08/891,227
/ PRIOR FILING DATE: 1997-06-24
/ NUMBER OF SEQ ID NOS: 368
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 102
/ LENGTH: 284
/ TYPE: PRT
/ ORGANISM: Helicobacter pylori
US-09-895-913A-102

Alignment Scores:
Pred. No.: 2,76e-18 Length: 284
Score: 246.00 Matches: 49
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.00% Mismatches: 0
Query Match: 15.78% Indels: 0
DB: 9 Gaps: 0

US-09-647-661-1 (1-900) x US-09-895-913A-102 (1-284)
QY 2 GCAGAAATTTACAGAGCCAGATAGCCAAATGGGGTTTATTTTAAATTTGAGCATG 61
Db 235 AlaGluPheTyArgGluProAspSerGlnMetGlyValTyPheLeuAsnLeuSerMet 254
QY 62 GGCACATTTTAAGCTTTATTTAGTGTGTTTCAATAGGATTTTATGATGCTACA 121
Db 255 GlyGlnIleLeuSerLeuPheMetValIleValSerLeuGlyIleLeuLeuTyAlaThr 274
QY 122 AAAAAATCTAAAAATAAAGGAAATCAA 151
Db 275 LysAsnSerLysLysIleLysGluAsnGln 284

RESULT 9
US-10-335-977-5526
/ Sequence 5526, Application US/10335977
/ Publication No. US20040052799A1
/ GENERAL INFORMATION:
/ APPLICANT: DOUGLAS SMITH et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
/ RELATING TO HELICOBACTER PYLORI FOR
/ DIAGNOSTICS AND THERAPEUTICS
/ NUMBER OF SEQUENCES: 10031
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: LAHIVE & COCKFIELD
/ STREET: 28 State Street
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02109-1875
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD-ROM ISO9660
/ OPERATING SYSTEM: Windows NT 4.0
/ SOFTWARE: UNIX
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/335,977
/ FILING DATE: 30-Dec-2002
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/993,002
/ FILING DATE: 17-DEC-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mandragouras, Amy E.
/ REGISTRATION NUMBER: 36,207
/ REFERENCE/DOCKET NUMBER: GTN-018
/ TELECOMMUNICATION INFORMATION:
```

```
/ TELEPHONE: (617)227-7400
/ TELEFAX: (617)742-4214
/ INFORMATION FOR SEQ ID NO: 5526:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 283 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEetical: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Helicobacter pylori
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (B) LOCATION 1...283
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5526:
US-10-335-977-5526

Alignment Scores:
Pred. No.: 2.18e-16 Length: 283
Score: 229.00 Matches: 45
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 93.75% Mismatches: 0
Query Match: 14.69% Indels: 0
DB: 12 Gaps: 0

US-09-647-661-1 (1-900) x US-10-335-977-5526 (1-283)
QY 2 GCAGAAATTTACAGAGCCAGATAGCCAAATGGGGTTTATTTTAAATTTGAGCATG 61
Db 235 AlaGluPheTyArgGluProAspSerGlnLeuGlyValTyPheLeuAsnLeuSerMet 254
QY 62 GGCACATTTTAAGCTTTATTTAGTGTGTTTCAATAGGATTTTATGATGCTACA 121
Db 255 GlyGlnIleLeuSerValPheMetValIleValSerLeuGlyIleLeuLeuTyAlaThr 274
QY 122 AAAAAATCTAAAAATAAAGGAA 145
Db 275 LysAsnSerLysLysIleLysGlu 282

RESULT 10
US-10-364-397-17
/ Sequence 17, Application US/10364397
/ Publication No. US20040014191A1
/ GENERAL INFORMATION:
/ APPLICANT: Minton, Nigel
/ APPLICANT: Anlezark, Gill
/ APPLICANT: Vaughan, Thomas
/ TITLE OF INVENTION: Nitroreductase Enzymes
/ FILE REFERENCE: 1581.0850001
/ CURRENT APPLICATION NUMBER: US/10/364,397
/ CURRENT FILING DATE: 2003-02-12
/ PRIOR APPLICATION NUMBER: 09/913,068
/ PRIOR FILING DATE: 2001-08-09
/ PRIOR APPLICATION NUMBER: PCT/GB00/00431
/ PRIOR FILING DATE: 2000-02-10
/ PRIOR APPLICATION NUMBER: UK 9903019.9
/ PRIOR FILING DATE: 1999-02-10
/ NUMBER OF SEQ ID NOS: 29
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 17
/ LENGTH: 220
/ TYPE: PRT
/ ORGANISM: Haemophilus influenzae
US-10-364-397-17

Alignment Scores:
Pred. No.: 3.01e-15 Length: 220
Score: 218.50 Matches: 59
Percent Similarity: 48.89% Conservative: 51
Best Local Similarity: 26.22% Mismatches: 94
Query Match: 14.02% Indels: 21
DB: 15 Gaps: 4
```

```

US-09-647-661-1 (1-900) x US-10-364-397-17 (1-220)
QY 151 ATGAATTTTGGATCAGAAAAGAGACAACTATTAAACGAGCGCCATTCTTGCAG 210
Db 1 MetThrGlnLeuThrArgGlnGlnValLeuGluLeuPheHisGlnArgSerThrArg 20
QY 211 ATGTTTGACGCCATTATAGTTTCTAGTGAAGAATAGAAAGAAATCGCTGAATCGCC 270
Db 21 TyrTyrAspProThrLysLysIleSerAspGluAspPheGluCysIleLeuGluCysGly 40
QY 271 AGGCTATCGCAAGCTCTTACACAGCGAGCGATGCGATTCTTGTGATGCTTACTAATAAG 330
Db 41 ArgLeuSerProSerValGlySerGluProThrLysPheLeuValIleGlnAsnLys 60
QY 331 GATTAAAAAACAATTCAGTCAGTCACAGCTACTTTAATGAAGAAATGATTAAAGCGCT 390
Db 61 ThrLeuArgGluLysMetLysProPheSerTrpGlyMetIleAsnGlnLeuAspAsnCys 80
QY 391 TCAGGTTAATGGGTATGCTTTAAGACCTAGCGAGTTGTACACAGCGGCATATAC 450
Db 81 SerHisLeuValValIleLeuAlaLysLys 94
QY 451 ATGCAAAACCTTTACCGGAGCTCTTATAAGGTAGAGTGATCCCTTCTTTGCTCAATG 510
Db 95 AspSerProPhePheValAspValMetAlaArgLysGlyLeuAsnAlaGluGlnGln 114
QY 511 CTGGCGTGAGATTCACACAGCATGCAAGATAGAAAGCTATATTTAGAG----- 564
Db 115 AlaAlaLeuThrLysTyrLysAlaLeuGlnGluAspMetLysLeuLeuGluAsnAsp 134
QY 565 -----CAATGCTATATCGCTGTGGGCAATTTGCGATG 597
Db 135 ArgThrLeuPheAspTrpCysSerLysGlnThrTyrIleAlaLeuAlaAsnMetLeuThr 154
QY 598 GCGGTGAGCTTAATCGGATGGATAGTTGCATTATGAGGCTTTGATCCCTTTAAAAAGTG 657
Db 155 GlyAlaSerAlaLeuGlyIleAspSerCysProIleGluGlyPheHisTyrAspLysMet 174
QY 658 GGTGAGTTTAA--GAGAGCGTATCAATAGCTTAA-----ATCGCATGCTTG 705
Db 175 AsnGluCysLeuAlaGluGluGlyLeuPheAspProGlnGluTyrAlaValSerValala 194
QY 706 ATCGCTTTGGGCAAGAGGTGGCAGAGCGAGCGCAAAATCAAGAAATCAAAAGTTGAT 765
Db 195 AlaThrPheGlyTyrArgSerArgAspIleAlaLysLysSerArgLysGlyLeuAspGlu 214
QY 766 GCGATTACTTGTTG 780
Db 215 ValValLysTrpVal 219

RESULT 11
US-10-364-397-23
; Sequence 23, Application US/10364397
; Publication No. US20040014191A1
; GENERAL INFORMATION:
; APPLICANT: Minton, Nigel
; APPLICANT: Anlezark, Gill
; APPLICANT: Vaughan, Thomas
; TITLE OF INVENTION: Nitroreductase Enzymes
; FILE REFERENCE: 1581.0850001
; CURRENT APPLICATION NUMBER: US/10/364,397
; CURRENT FILING DATE: 2003-02-12
; PRIOR FILING DATE: 09/913,068
; PRIOR APPLICATION NUMBER: PCT/GB00/00431
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: UK 9903019.9
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 201
; TYPE: PRT

US-09-647-661-1 (1-900) x US-10-364-397-23 (1-201)
QY 169 GAAAAAGAGAACAACTATTAAACGAGCGCCATTCTTGAAGATGTTGACAGCCATTAT 228
Db 2 LysLysGluLeuGluIlePheSerThrArgTyrSerCysArgAsnPheLysAsnGlu--- 20
QY 229 GAGTTTCTAGTGAAGAATAGAAAGAAATCGCTGAATCGCAGCGCTATCGCAAGCTCT 288
Db 21 LysLeuLysGluAspLeuAsnSerIleLeuGluIleAlaArgLeuSerProSerSer 40
QY 289 TACAACACGACGATGCGATTGTTGATGTTTACTAATAGGATTTAAAAAACAAT 348
Db 41 LeuGlyLeuGluProTrpLysPheIleValValGlnAspGluLysArgLysGluLeu 60
QY 349 GCGAGTCACAGCTACTTTAATGAAGAAATG---ATTAAAGCGCTTCACGTTATGNG 405
Db 61 Ser-----LysIleCysAsnGlnGlnLysHisValLysAspCysAlaAlaLeuIleIle 78
QY 406 GTATGCTCT-----TTAAGACCTAGCGAGTTG 432
Db 79 IleIleSerArgLeuAspPheLeuAspTyrPheGluGluLysLeuArgLysArgAspMet 98
QY 433 TTACACAGCGCCATTACATGCAAAACCTTTACCGGAGCTCTTATAGGTAGATGATC 492
Db 99 -----SerGluThrGluMetGlnLysArgLeu 107
QY 493 CCTCTTTGCTCAATGCTTGGCGTGAGATTCAACACAGCATGCAAAAGATTAGAAAGC 552
Db 108 AspThrTyrMetProPheLeuLys-----SerLeuAsnGlnGluGlnLysIleSer 124
QY 553 TATATTTTGAAGCAATGCTATATCGCTGTGGGCAAAATTTGATGGCGGTAGCTTAATG 612
Db 125 TyrAlaArgGluGlnAlaHisIleAlaLeuAlaSerIleLeuTyrSerAlaAsnAlaLeu 144
QY 613 GGATTCGATAGTTGCTATTTGAGGCTTTGATCCCTTTAAAGTGGGTCAAGTTTAGAA 672
Db 145 AsnIleAlaSerCysThrIleGlyGlyPheAspLysGluLysLeuAspSerTyrLeuSer 164
QY 673 GAGCGTATCAATAAGCCTAAATCGCATGCTTATCGCTTTGGGC 717
Db 165 LeuAspIleGlnLysGluArgSerSerLeuValValAlaLeuGly 179

RESULT 12
US-10-335-977-6413
; Sequence 6413, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1975
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible

```

OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-1214
INFORMATION FOR SEQ ID NO: 6413:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCAT 1...217
SEQUENCE DESCRIPTION: SEQ ID NO: 6413:
US-10-335-977-6413
Alignment Scores:
Pred. No.: 4,54e-12 Length: 217
Score: 130.00 Matches: 54
Percent Similarity: 51.63% Conservative: 57
Best Local Similarity: 25.12% Mismatches: 86
Query Match: 12.15% Indels: 18
DB: 7 Gaps: 7
US-09-647-661-1 (1-900) x US-10-335-977-6413 (1-217)
QY 160 TTGGATCAGAAAAAGAGACAACTATTAAACGAGCGCATTTTCAAGATGTTTGCAC 219
Db 1 MetAspArgGluGlnIleleAlaLeuGlnHisGlnArgPheAlaThrLysLysTyrAsp 20
QY 220 AGCCATTATGAGTTTCTAGTGAAGAATAGAAATGCTGAAATGCCAGCGGTATCG 279
Db 21 ProAsnArgArgIleSerGluLysAspTrpGluValLeuValGluValGlyArgLeuAla 40
QY 280 CCAAGCTTTCAACACGAGCGCATTTTGTGATGTTTACTAATAAGGATTTAAA 339
Db 41 ProSerIleGlyLeuGluProIrpLysMetLeuLeuLysAsnGluArgMetLys 60
QY 340 AAACAAATTCAGTCACAGCTACTTTAATGAAGAAATGATTAAAGCGCTTCAGCGTTA 399
Db 61 GluAspLeuLysProMetAlaTrpGlyGlyLeuSerSerLeuGluGlyAlaSerHisPhe 80
QY 400 ATGGTGGTATGCTTTAGACCTAGCGAGTTGTTACACAGCGCCATACATGCAAAAC 459
Db 81 ValIleTyr-----LeuAlaArgLysGlyValThrTyrAspSerAspTyrValLysLys 98
QY 460 CTTTACCCGAGTCTTATAAG-----GTAGAGTGATC 492
Db 99 ValMetHisGluValLysLysArgAspTyrAspThrHisSerArgPheAlaGlnIlele 118
QY 493 CCTTCTTTTGCATTAATGCTGGGTGAGATTCAACACAGCATGCAAGATTAAGAGC 552
Db 119 LysAsnPhe---GlnGluAsnAspIleLysLeuAsn---SerGluArgSerLeuPheAsp 136
QY 553 TATATTTTAGAGCAATGCTATATCGCTGTGGGGCAAAATTTGCTGCGGTGAGCTTAATG 612
Db 137 TrpAlaSerLysGlnThrTyrIleGlnMetAlaAsnMetMetMetAlaAlaMetLeu 156
QY 613 GGATTGGATGTTGCATTATTCGAGCTTTGATCCCTTTAAAGAGGGGTGAAGTTTGA 672

Db 157 GlyIleAspSerCysProIleGluGlyTyrAspGlnGluLysValGluAlaTyrLeuGlu 176
QY 673 GAGCGT-----ATCAATAGCCTAAA-----ATCGCATGCTTGATCGCTTTGGCGAAG 720
Db 177 GluLysGlyTyrLeuAsnThrAlaGluPheGlyValSerValMetAlaSerPheGlyTyr 196
QY 721 AGGGTGGCAGACGAGCGCCAAATAATCAAGA---AAATCAAAAGTT 762
Db 197 ArgAsnGlnGluIleThrProLysThrArgTyrLysThrGluVal 211
RESULT 13
US-10-364-397-12
; Sequence 12, Application US/10364397
; Publication No. US20040014191A1
; GENERAL INFORMATION:
; APPLICANT: Minton, Nigel
; APPLICANT: Anlezark, Gill
; APPLICANT: Vaughan, Thomas
; TITLE OF INVENTION: Nitroreductase Enzymes
; FILE REFERENCE: 1581.0850001
; CURRENT APPLICATION NUMBER: US/10/364,397
; CURRENT FILING DATE: 2003-02-12
; PRIOR APPLICATION NUMBER: 09/913,068
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: PCT/GB00/00431
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: UK 9903019.9
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-364-397-12
Alignment Scores:
Pred. No.: 1.41e-11 Length: 202
Score: 185.50 Matches: 53
Percent Similarity: 49.76% Conservative: 50
Best Local Similarity: 25.60% Mismatches: 85
Query Match: 11.90% Indels: 19
DB: 7 Gaps: 7
US-09-647-661-1 (1-900) x US-10-364-397-12 (1-202)
QY 184 CTATTAAACGAGCGCCATTTTCAAGATGTTTGCAGCGCATTTAGTATGATTTTCTAGTAA 243
Db 7 ValLeuLysAlaArgAlaSerValLysGluTyrAspThrAsnAlaProIleSerLysGlu 26
QY 244 GAATTAGAGAAATCGCTGAAATCGCAGCGCTATCGCCAAAGCTCTTTACACACGCGCA 303
Db 27 GluLeuThrGluLeuLeuAspLeuAlaThrLysAlaProSerAlaTrpAsnLeuGlnHis 46
QY 304 TGGCATTTTGTGATGTTACTAATAAGGATTTAAATAAAACAATTCAGTCGACGACTAC 363
Db 47 TrpHisPheThrValPheHisSerAspGlySerLysAlaGluLeuLeuProValAlaTyr 66
QY 364 TTTAATGAAGAAATGATTAAAGCGTTTCAAGCGTTTAAATGTTGATGTTTAAAGACCT 423
Db 67 AsnGlnLysGlnIleValGluSerSerAlaValAlaIleLeuLysAspLeuLysAla 86
QY 424 AGCGAGTTGTTACACAGCGCCATTTACATGCAAAACCTTTACCCGAG---TCTTTAAG 480
Db 87 AsnGlu-----AsnGly-----GluGluValTyrAlaGluLeuAlaSerGln 100
QY 481 GTTAGAGTGATCCCTTCTTTTGTCTCAATGCTTGGGTGAGATTCACACACGACGACAA 540
Db 101 GlyTyrIleThrAspGluLysGlnThrLeuLeuGlyGlnIleAsnGlyAlaTyrGln 120
QY 541 AGATTGAAGAGCTATATTTTAGAGCAATGCTAT-----ATCGTGTGGGGCAA 588

Db 121 Ser---GluGlnPheAlaArgAspSerAlaPheLeuAsnAlaSerLeuAlaAlaMetGln 139
QY 589 ATTTCATGGCGTGAAGCTTAATGGGATTGGATAGTTCATTAATGGAGGCTTTCATCCT 648
Db 140 LeuMetIleAlaAlaLysAlaLysGlyTyrAspThrCysAlaIleGlyGlyPheAsnLys 159
QY 649 -----TAAAGTGGCTGAAGTTTGTAGAGAGCGGTATCAATAAGCCTAAATCGCA 699
Db 160 GluGlnPheGlnLysGlnPheAspIleSerGluA-gTyrValPro-----Val 175
QY 700 TCGTTGATCGCTTTGGCGAAGGGTGGCAGAGCGAGCCAAATCAAGAAATCAAAA 759
Db 176 MetLeuIleSerIleGlyLysAlaValLysProAlaHisGlnSerAsnArgLeuProLeu 195
QY 760 GTTGATGCGATTAATTGGTTG 780
Db 196 SerLysValSerThrTrpLeu 202
RESULT 14
US-10-169-048-2
; Sequence 2, Application US/10169048
; Publication No. US20030072769A1
; GENERAL INFORMATION:
; APPLICANT: Clarke, Edna Elizabeth
; APPLICANT: Zhou, Liling
; APPLICANT: Shea, Jacqueline Elizabeth
; APPLICANT: Feldman, Robert Graham
; APPLICANT: Holden, David William
; TITLE OF INVENTION: Streptococcus Pyogenes Virulence Genes and Proteins And Their Use
; FILE REFERENCE: GUE-97
; CURRENT APPLICATION NUMBER: US/10/169,048
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: PCT/GB00/04997
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 221
; ORGANISM: Streptococcus pyogenes
US-10-169-048-2
Alignment Scores:
Pred. No.: 3 57e-11 Length: 221
Score: 182.00 Matches: 58
Percent Similarity: 47.06% Conservative: 46
Best Local Similarity: 26.24% Mismatches: 83
Query Match: 11.67% Indels: 34
DB: 14 Gaps: 6
US-09-647-661-1 (1-900) x US-10-169-048-2 (1-221)
QY 178 AGACAATATAACGAGCGCCATCTTCGCAAGATGTTGACAGCCATTATGAGTTTCT 237
Db 10 GlnGlnAlaLeuHisPheArgThrAlaValArgValTyrLysGluGlu---LysIleSer 28
QY 238 AGTGAAGAATTAGAGAAGAAATCGTGAATCGCAGGTCATCGCAAGCTCTTACACACG 297
Db 29 AspGluAspLeuAlaLeuIleLeuAspAlaAlaTrpLeuSerProSerSerIleGlyLeu 48
QY 298 CAGCCATGGCATTTGTGATGGTTACTATAAGGATTTAAAAAACAATAATGCGAGTGCAC 357
Db 49 GluGlyTrpArgPheValValLeuAspAsnLysProIleLysGluGluLysProPhe 68
QY 358 AGCTACTTTAATGAGAAGATTAAGAGCGCTCAGCGTTATGTTGGTATGCTCTTTA 417
Db 69 AlaTrpGlyAlaGlnTyrGlnLeuGluThrAlaSerHisPheIleLeuLeuIleAlaGlu 88
QY 418 AGACCTAGCGAGTTGTTTACCACACGCGCCATTACATGCAAAACCTTTACCGGAGTCTTAT 477
Db 89 Lys-----HisAlaArgTyrAspSerProAlaIleLysAsnSerLeu 102
QY 478 AAGTTAGATGATCCCTCTTTTGTGCTCAATGCTTGGCGTGAGATTCACACAGCATG 537

Db 103 LeuArgArgGlyIleLysGlu-----GlyAspGlyLeuAsnSerArgLeu 117
QY 538 CAAAGATTGAAGAGCTATATTTAGAG----- 564
Db 118 LysLeuTyrGluSerPheGlnLysGluAspMetAspMetAlaAspAsnProArgAlaLeu 137
QY 565 -----CAATGCTATATCGCTGTGGGCAAAATTTTCATGGCGGTGAGC 606
Db 138 PheAspTrpThrAlaLysGlnThrTyrIleAlaLeuGlyAsnMetMetThrAlaAla 157
QY 607 TTAATGGGATGATAGTTCATTAATGGAGGCTTTCATCCTTTAAAGTGGTGAAGTT 666
Db 158 LeuLeuGlyIleAspThrCysProIleGluGlyPheHisTyrAspLysValAsnHisIle 177
QY 667 TTGAAGACAGCGT-----ATCAATAAGCTTAAATCGCATGCTTGCATCGCTTTG 714
Db 178 LeuAlaLysHisAsnValIleAspLeuGluLysGluGlyIleAlaSerMetLeuSerLeu 197
QY 715 GCGAAGAGGGTGGCAGAGCGAGCCAA---AAATCAAGAAATCAAAAGTTGATGCGATT 771
Db 198 GlyTyrArgLeuArgAspProLysHisAlaGlnValArgLysProLysGluGluValIle 217
QY 772 ACT 774
Db 218 Ser 218
RESULT 15
US-10-364-397-19
; Sequence 19, Application US/10364397
; Publication No. US20040014191A1
; GENERAL INFORMATION:
; APPLICANT: Minton, Nigel
; APPLICANT: Anlezark, Gill
; APPLICANT: Vaughan, Thomas
; TITLE OF INVENTION: Nitroreductase Enzymes
; FILE REFERENCE: 1581.0850001
; CURRENT APPLICATION NUMBER: US/10/364,397
; CURRENT FILING DATE: 2003-02-12
; PRIOR APPLICATION NUMBER: 09/913,068
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: PCT/GB00/00431
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: UK 9903019.9
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 19
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Synechocystis PCC6803
US-10-364-397-19
Alignment Scores:
Pred. No.: 2 71e-10 Length: 200
Score: 174.00 Matches: 53
Percent Similarity: 47.50% Conservative: 42
Best Local Similarity: 26.50% Mismatches: 87
Query Match: 11.16% Indels: 18
DB: 15 Gaps: 6
US-09-647-661-1 (1-900) x US-10-364-397-19 (1-200)
QY 193 GAGCGCCATCTTCGCAAGATGTTGACAGCCATTATGAGTTTCTAGTGAAGAATTAGAA 252
Db 9 GlnAlaArgSerValLysHisPheAspProAspHisArgLeuThrAlaGluGluArg 28
QY 253 GAAATCGGTGAATCGCAGGCTATCGCAAGCTCTTACACACGACGCCATGGCATTTT 312
Db 29 LysLeuHisGluAlaAlaIleGlnAlaProThrSerPheAsnIleGlnLeuTrpArgPhe 48
QY 313 GTGATGGTTACTTAATAAGGATTTAAAAAACAATAATGCGATGTCAGCATCTTTAATGAA 372

```
Db 49 LeuileileArgaspProGlnLeuArgGlnThrIleArg---GluLysTyrGlyAenGln 67
QY 373 GAAATGATTAAAGCGCTTCAGCGTTAATGGTGTATGCTCTTTAAGACCTAGCGAGTTG 432
Db 68 AlaGlnMetThrAspAlaSerLeuLeuileuValAlaAlaaspValasnAlaIrtPasp 87
QY 433 TTACCACACGCGCATACATGCAAAACCTTTACCGGAG-----TCTTATAAGGTTAGA 486
Db 88 LysAspProAlaArgTyrTrpArgAsnAlaProArgGluValAlaAlaAsnTyrLeuValGly 107
QY 487 GTGATCCCTTCTTTTGGCTCAATGCTTGGCGTGAGATTCAACCACAGCATGCAAGATT 546
Db 108 AlaileAlaPheTyrGlyGly-----LysProGlnLeuGlnArgasp 121
QY 547 GAAAGCTATATTTAGAGCAATGCTATATCGTGTGGGCAAAATTTGCATGGCGGTGAGC 606
Db 122 GluAlaGln-----ArgserIleGlyMetAlaMetGlnAsnLeuMetLeuAlaAlaLys 139
QY 607 TTAATGGGATTGGATGTCATTATTTGGAGGCTTTGATCCTTTAAAGTGGGTGAGTT 666
Db 140 AlaMetGlyTyrAspSerCysProMetIleGlyPheAspLeuGlnLysValAlaGluLeu 159
QY 667 TTAGAAGAGCGTATCATAAGCCT-----AAAATCGCATGCTTGATCGCTTTGGGC 717
Db 160 -----ValLysLeuProAlaAspTyrAlaIleGlyProMetValAlaIleGly 175
QY 718 AAGAGGTGGCAGACGACCCCAAAATCAAGAAATCAAAAGTTGATGCGATTACTTGG 777
Db 176 LysArgThrGluaspAlaArgAlaLysGlyGlyGlnThrProLeuGluGluLeuValTrp 195
```

Search completed: April 20, 2004, 19:22:36
Job time : 68 secs